THE UNIVERSITY
of ADELAIDE

Edepillim: A New Muon Energy Reconstruction for the IceCube Neutrino Observatory

Sally Alice Robertson

B.Sc (Honours)

A thesis submitted to the University of Adelaide
in fulfilment of the requirements for the degree
of Doctor of Philosophy.

School of Physical Sciences

Department of Physics

May 2018

Contents

Abstract

Neutrinos are an ideal tracer particle for high energy astrophysics as they are not deflected by magnetic fields. As such they can be detected at Earth with their directional information intact, however because of the low interaction probability they are difficult to detect. The IceCube Neutrino Observatory operates by observing Cherenkov radiation emitted from charged secondary interaction particles of neutrinos such as muons. As a muon travels though the detector it experiences energy loss due to radiative processes which will be represented by an energy loss pattern with stochastic losses, the probability of these losses is dependent on the initial energy of the muon.

The muon energy reconstruction method Edepillim described in this work takes into account the entire pattern of energy loss along the muon track and uses probability distribution functions describing muon energy losses to perform a maximum likelihood reconstruction for the initial muon energy. This work demonstrates the good reconstruction resolution of this method on idealised simulation by comparison to other energy reconstruction methods. In addition the method was implemented on simulated muon events where the pattern of energy losses have been reconstructed, and the limitations and improvements needed to be made to the energy loss pattern for optimal resolution are discussed. Finally in this work the energy reconstruction method was tested on a sample of up-going neutrino-induced muon events, by performing a diffuse fit to atmospheric and astrophysical flux predictions. This fit demonstrates the improvement that this energy resolution has on the confidence interval on the value of the fitted astrophysical flux results. Also discussed is a specific example of reconstructing IceCube's highest energy muon neutrino event using Edepillim.

Declaration of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Acknowledgements

I would like to thank my supervisors Gary Hill and Bruce Dawson for their support and guidance throughout my PhD, and for giving me the opportunities to grow and learn in this process. I would particular like to thank Gary for his tremendous support throughout the writing process and the late nights he dedicated to assisting me. I would also like to thank Ben Whelan for his teaching, technical support and advice.

Also, thank you to my fellow members of the High Energy Astrophysics Group in Adelaide who over the years have made this a friendly environment to work in. A special thank you to my fellow students who have worked with me and sympathised with me; Mark, Alex W, Alex K, Natasha and Rebecca.

I would also like to thank all the IceCube collaboration members, who have worked to make the collaboration an exciting and encouraging group to work within, I would particular like to thank Kyle Jero and Leif Radel for their assistance in my own work.

I would like to thank my family who have always supported me and encouraged me in all my purists. My parents Sue and Stuart who have made it possible for me to achieve my dreams and never stopped me from following my passion. Thank you to the rest of my family Jayne, Claire, Andrew and Nicholas who have always been there for me with encouragement and hugs when I needed it. Thank you to my partner Matthew who has joined me on this journey and supported me even when he had his own thesis to write.

Finally I would like to acknowledge all my friends and family whose love and support got me to where I am today.

Authors Note

It is important to note the contributions from other individuals that impacted this work and to highlight the contributions that were made by the author. This work was performed as a member of the IceCube collaboration and used some resources produced by the collaboration.

The background material relevant to this thesis is presented in Chapters 1 to 3. With Chapter 1 summarising the previous knowledge of cosmic rays and neutrinos, Chapter 2 introducing the IceCube neutrino observatory and its current results, and Chapter 3 summarising the background knowledge of muons as well as the reconstruction and simulation methods used in IceCube for muons.

The early development of the Edepillim energy reconstruction method described in this work was performed by Gary Hill and Mark Aartsen before the author joined the group. The work in Chapters 4, 5, 6, 7 and 8 are the key chapters of this thesis and are all the work of the author.

The simulations used in this work are summarised in Appendix A. The work in Chapters 6, 7 and 8 uses simulations produced by the IceCube collaboration.

Chapter 1

Introduction: High Energy Astrophysics

Astrophysical sources are believed to produce high energy cosmic rays, neutrinos and gamma rays. In this chapter we discuss the acceleration of cosmic rays as well as the production of gamma rays and neutrinos. We also discuss the expected energy spectrum for astrophysical sources for such events and how neutrinos can be used as tracer particles for detection of the source spectrum and properties of the source spectrum.

1.1 High Energy Cosmic Rays

Cosmic Rays are charged particles produced and accelerated by various sources in the Universe. Victor Hess discovered cosmic rays in 1912 when his balloon experiments measured ionising radiation hitting the atmosphere [1]. Since their discovery, multiple experiments have measured their energy spectrum and the composition of the cosmic rays arriving at Earth. The energy spectrum in Figure 1.1 shows the results from various experiments that detect cosmic rays at different energy ranges. The overall spectrum approximately follows a power law,

$$\frac{dN}{dE} \propto E^{-\gamma} \tag{1.1}$$

where dN/dE is the differential flux of cosmic rays and γ is the spectral index that the flux follows. The spectral index γ has been found not to be constant across all cosmic ray energies. There are two distinct features that define a change in the value of γ known as the ‘knee’ and the ‘ankle’. Cosmic rays are thought to come from various extraterrestrial sources including galactic and extragalactic sources. It is postulated that the feature known as the ‘knee’ of the energy spectrum near 10^{15} eV (see Figure 1.1) is caused by the decrease in galactic cosmic rays as 3×10^{15} eV is the maximum of the energy of particles accelerated in supernova shocks [2]. At energies below this feature the cosmic ray flux follows the spectral index $\gamma = 2.7$, at energies above the knee the spectral index is $\gamma = 3.0$.

The flattening of the spectrum known as the ‘ankle’ is believed to be a transition from galactic sources to extragalactic sources [2]. Possible extragalactic sources include active galactic nuclei and gamma ray bursts [3].

1.2 Acceleration of High Energy Cosmic Rays

Cosmic rays are thought to be accelerated due to encounters with magnetic fields which will accelerate particles but will also keep them contained. The size of the source needs to be compatible with the Larmor radius for the particle to be contained. The Larmor radius in parsecs is given by

$$r_L(pc) = 1.08 \frac{E(\mathrm{PeV})}{ZB(\mu\mathrm{G})} \tag{1.2}$$

where Z is the charge of the particle, E is the particle’s energy in units of PeV and B is the component of the magnetic field normal to the particle’s velocity in microgauss [5]. As the acceleration causes an increase in the particle’s energy within the source, the radius required to contain the particle increases. This causes the particle to eventually escape the source region and have no further increase in its energy.

The acceleration of particles via interaction with a magnetised interstellar medium was first proposed by Fermi [6]. Fermi’s original theory, now called 2nd order Fermi acceleration, involves the collision of a particle with a cloud of interstellar medium, transferring the energy of the macroscopic motion of the cloud to the microscopic particle through magnetic inhomogeneities. In the example of interstellar clouds, random velocities can lead to an energy gain of the order of

$$\frac{\Delta E}{E} \propto \left(\frac{v}{c}\right)^2 \tag{1.3}$$

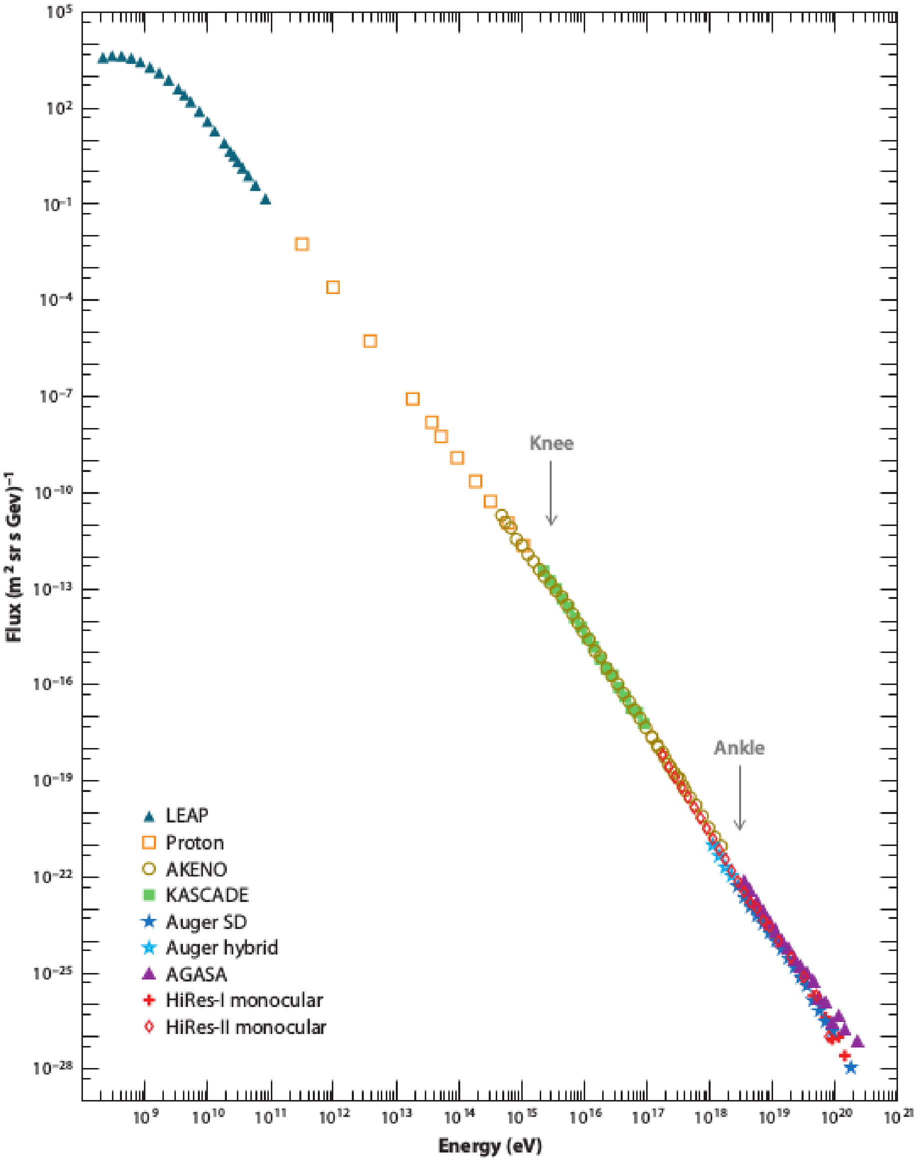


Figure 1.1: The cosmic ray energy spectrum as measured by multiple experiments, the breaks in the spectrum known as the knee and ankle are indicated. Figure from [4].

where E is the initial energy of the particle, ΔE is the change in energy and v is the speed of the interstellar cloud [3].

This process is not effective at accelerating high energy cosmic rays, however by applying the process to a shock front, it is found that a particle can bounce back and forth multiple times. After n encounters with the shock front the particle has acquired energy

$$E_n = E_0(1 + \varepsilon)^n \tag{1.4}$$

where E_n is the final energy, E_0 is the initial energy and $\varepsilon = \Delta E/E$ is the average relative energy gain. The number of encounters with the shock front to reach energy E_n is given by

$$n = \ln\left(\frac{E_n}{E_0}\right) / \ln(1 + \varepsilon) \tag{1.5}$$

For each encounter there is a probability, P_{esc}, that the particle will escape and no longer be accelerated. The probability that a number of particles N can reach an energy greater than E can be given by

$$N(> E) \propto \sum_{m=n}^{\infty} (1 - P_{esc})^m \tag{1.6}$$

The summation term in Eq 1.6 can be simplified such that

$$N(> E) \propto \frac{(1 - P_{esc})^n}{P_{esc}} \tag{1.7}$$

which can be combined with Eq 1.5 to form the integral energy spectrum

$$N(> E) \propto \frac{1}{P_{esc}} \left(\frac{E}{E_0}\right)^{\alpha} \tag{1.8}$$

where

$$\alpha = \ln\left(\frac{1}{1 - P_{esc}}\right) / \ln(1 + \varepsilon) \tag{1.9}$$

This describes an energy spectrum in terms of a power law that is consistent with the observed flux at Earth as seen in Figure 1.1.

The spectral index α can be simplified to $\alpha \sim P_{esc}/\varepsilon$ via the assumptions that the escape probability is very low and the energy gain per encounter is very small (i.e. $\varepsilon << 1$) [7]. The value for α depends on the ratio of shocked to unshocked gas velocities. As calculated in [8] the spectral index can be approximated by $\alpha \sim 1$. For the case of the differential energy spectrum the spectral index is given by $dN/dE \propto E^{-\gamma}$ where $\gamma = \alpha + 1$, thus predicting a source spectrum of $dN/dE \propto E^{-2}$. The final energy spectrum at Earth will also include propagation effects for the cosmic rays in the galaxy. The effect of propagation will add an energy dependence of $E^{-0.6}$ [8] which will result in a spectral index at Earth of $\gamma = 2.6$, which is in good agreement with the observed values of $\gamma = 2.7$ for energies below the knee [9].

1.3 Astrophysical Neutrino and Gamma Ray Production

After the acceleration of cosmic rays at the source, cosmic rays can interact with the surrounding medium to produce gamma rays and neutrinos (See Section 1.5). Theoretical astrophysical shock fronts where cosmic rays are believed to be produced would result with accelerated comic rays, most likely protons, interacting with the surrounding medium such as protons, neutrons or photons. This is demonstrated in Figure 1.2. The interactions with matter and photons are as follows:

$$p + p \rightarrow \begin{cases} p + p + \pi^+, & \text{fraction } 2/3 \\ p + n + \pi^-, & \text{fraction } 1/3 \end{cases} \tag{1.10}$$

$$p + \gamma \rightarrow \triangle^+ \rightarrow \begin{cases} p + \pi^0, & \text{fraction } 2/3 \\ n + \pi^+, & \text{fraction } 1/3 \end{cases} \tag{1.11}$$

The resulting pions then decay, with neutral pions decaying as follows

$$\pi^0 \rightarrow \gamma + \gamma \tag{1.12}$$

and producing gamma rays [10]. This results in another astrophysical tracer, where cosmic rays, neutrinos and gamma rays should all be produced at the same source as shown in Figure 1.2.

The charged pions will decay to muons and neutrinos

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{1.13}$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{1.14}$$

The charged muons will also decay and increase the amount of neutrinos produced

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{1.15}$$

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \tag{1.16}$$

the result being that in high energy astrophysical sources you expect to have the production of cosmic rays, gamma rays and neutrinos. Gamma rays are further produced at the source by the charged cosmic ray particles themselves, any charged particle moving at angle to a magnetic field will lose energy to synchrotron radiation with the production of gamma ray photons. Most synchrotron radiation is produced by electrons as the energy loss is inversely proportional to the square of the particle mass [8].

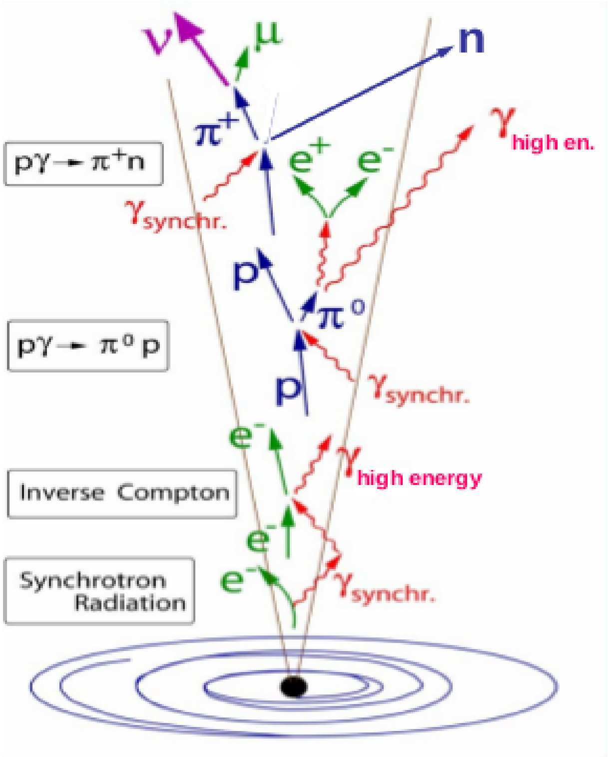


Figure 1.2: The production of cosmic rays, gamma rays and neutrinos at astrophysical shock fronts. Figure from [10].

If the production of neutrinos occurs immediately after the production of cosmic rays with no other processes causing loss of energy, it is expected that the neutrino energy spectrum will follow that of Fermi shock acceleration with a spectrum of E^{-2}. This differs from the cosmic ray spectrum in Section 1.2 as that spectrum takes into account the propagation effects of charged particles in the galaxy [8].

1.4 Limits to Detection of High Energy Cosmic Rays

The difficulty with finding the sources of cosmic rays and thus understanding their acceleration mechanism lies in the nature of the particles. They are charged particles and so are deflected and delayed by magnetic fields during travel from their original source, although particles of ultra high energy (above $\sim 10^{18}$ eV) are expected to maintain information about their original direction.

Other difficulties in detection are due to the GZK effect wherein particles with energies above 6×10^{19} eV will interact with the cosmic microwave background radiation. Protons above this energy would lose energy by the following processes

$$\gamma_{CMB} + p \rightarrow \Delta^{+} \rightarrow n + \pi^{+} \tag{1.17}$$

$$\gamma_{CMB} + p \rightarrow \Delta^{+} \rightarrow n + \pi^{0} \tag{1.18}$$

which causes a cut off in the spectrum and produces neutrons as well as positive and neutral pions [11]. The neutral pions in Eq 1.18 will decay into gamma rays and the charged pions in Eq 1.17 will decay into muons and muon neutrinos

$$\pi^{+} \rightarrow \mu^{+} + \nu_{\mu} \tag{1.19}$$

The deflection and absorption of cosmic rays limit their use as tracer particles to astrophysical acceleration sources.

1.5 Neutrinos

1.5.1 Background

The neutrino was first proposed in 1930 by Wolfgang Pauli in order to explain conservation of momentum and energy in Beta decay [12]. Beta decay occurs when there are an unstable number of protons and neutrons in a nucleus. The neutron decays to a proton and ejects daughter particles. Prior to Pauli's proposal it was thought that the daughter particle would be a single electron. However when the daughter products of beta decay were measured experimentally it was found that the energy and momentum had not apparently been conserved. Thus Pauli postulated that a weakly interacting neutral particle with spin $\frac{1}{2}$ was also produced.

In 1934 Fermi published a paper (translated here [13]) which described the mechanics of beta decay which also included Pauli's proposal of including a neutrino. The model of beta decay now included the ejection of an electron and a neutrino,

$$n \rightarrow p + e^- + \bar{\nu}_e \tag{1.20}$$

At the time it was thought the neutrino would be completely undetectable. However in 1956 Clyde Cowan and Frederick Reines were able to observe an anti-neutrino induced inversion of neutron decay, with the anti-neutrino originating from fission products in a nearby reactor [14],

$$\bar{\nu}_e + p \rightarrow e^+ + n \tag{1.21}$$

The reaction was observed via an anti-neutrino incident on a water target with Cadmium Chloride dissolved within the solution. The anti-neutrino produces a positron and neutron, and in a short time frame the positron annihilates with an electron

$$e^+ + e^- \rightarrow \gamma + \gamma \tag{1.22}$$

giving two 0.5 MeV annihilation gamma rays which were detected in prompt coincidence by scintillation detectors on opposite sides of the system. The neutron is captured by cadmium, releasing gamma rays, which shows a characteristic delayed coincident count with the subsequent annihilation gamma rays.

$$n + Cd \rightarrow \gamma + \gamma \tag{1.23}$$

The experiment detected the first pulse of positron annihilation and the second from the neutron capture. In addition the signal was dependent on the amount of target protons. The detection results in the conclusion of a neutrino present due to the fission products in Eq 1.21.

Since its first detection the neutrino has been studied to bring us to our current understanding, where neutrinos are part of the lepton family with three flavours; muon (ν_μ), electron (ν_e) and tau (ν_τ).

Neutrinos are theorised to be produced in high energy astrophysical objects, along with cosmic rays and gamma rays (see Section 1.3). Detection of these astrophysical objects is impaired by the absorption of gamma rays and the deflection of charged cosmic rays. However neutrinos are neutral particles that are not deflected by galactic or extragalactic magnetic fields thereby making it possible to find their point of origin, since they travel in straight lines. This makes neutrinos the ideal tracer particle to use in high energy particle astrophysics. Neutrinos interact only via the weak force and so have a high probability of surviving to Earth intact. At low energies the cross-section of interaction is low making them difficult to detect, however at higher energies the cross-section is higher allowing for detection but also resulting in neutrinos not reaching the detector, this is discussed more in Section 1.5.3.

1.5.2 Neutrino Oscillation

For the proton-proton interaction production of neutrinos as described in Section 1.3 the flavour ratio at the production source is $(\nu_\mu : \nu_e : \nu_\tau) = (2 : 1 : 0)$. However, due to the neutrinos having non-vanishing mass eigenstates the neutrinos will oscillate as they travel through space such that the flavour ratio becomes $(1 : 1 : 1)$ at Earth.

The two methods for oscillation are vacuum and matter oscillations. The vacuum oscillations involve the mixing of mass and flavour eigenstates. The second method is due to flavour dependent forward scattering amplitudes in matter, known as the MSW effect. Vacuum oscillations are the simpler case as the Hamiltonian remains constant [15]. However in matter the Hamiltonian is non-constant as properties of the medium can change along the neutrino path. In this section we describe the oscillation in a vacuum.

Neutrino flavour-eigenstates, denoted as $|\nu_\alpha\rangle$,with $\alpha = e, \mu, \tau$, are a superposition of the mass eigenstates, denoted as $|\nu_j\rangle$, with $j = 1, 2, 3$, via the following

$$|\nu_\alpha\rangle = \sum U_{\alpha j} \exp{(-iE_j t)} |\nu_j\rangle \tag{1.24}$$

where E_j is the eigenvalue for the Hamiltonian operator in a vacuum, therefore the energy of the mass-eigenstate, t is the time from emission at the source, $U_{\alpha j}$ is the mixing matrix which depends on three mixing angles θ_{12}, θ_{13}, θ_{23} and a Dirac phase δ. The mixing matrix for the assumption of three massive neutrinos is

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.25}$$

where $s_{jk} = \sin\theta_{jk}$, $c_{jk} = \cos\theta_{jk}$, $j = 1, 2$, and $k = 1, 2, 3$, where $j \neq k$.

The probability of a neutrino to oscillate from flavour state α to β is given by

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= |\langle \nu_\beta(t)|\nu_\alpha(t=0)\rangle|^2 \\ &= \delta_{\alpha\beta} - 4\sum_{j>i} U_{\alpha i}U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2\frac{\Delta m_{ij}^2 L}{4E_\nu} \end{aligned} \tag{1.26}$$

Here time is given as $t = 0$ to indicate the mixing at the neutrino source and Δm_{ij}^2 is the difference of the squared masses of two flavours i and j. [16]

The three mixing angles can be approximated as $\theta_{12} = \pi/6$, $\theta_{23} = \pi/4$, $\theta_{13} = 0$ [17] [18]. These values come from measurements of neutrino fluxes by different experiments. The measurement of oscillations from electron neutrinos to muon neutrinos, yielding θ_{12}, comes from solar neutrino detections such as Super-Kamiokande and the Sudbury Neutrino Observatory. Oscillation from muon neutrinos to tau neutrinos, yielding θ_{23}, can be measured from atmospheric neutrinos. The oscillations from election neutrinos to tau neutrinos, using the mixing parameter θ_{13} are the smallest and limits on this process come from reactor experiments, such as Double Chooz, Daya Bay and RENO.

For the above mixing angle approximations the mixing matrix can be given by

$$U = \begin{pmatrix} 0.866 & 0.5 & 0 \\ -0.354 & 0.612 & 0.707 \\ 0.354 & -0.612 & 0.707 \end{pmatrix} \tag{1.27}$$

For cosmological origin the probability is averaged over the energy spectrum of neutrinos, which leads to the simplification in Eq 1.26 of $\sin^2(\Delta m_{ij}^2 L/4E_\nu) \sim 0.5$ for large distances. Thus the oscillation probability is independent of time for cosmological distances.

Using these assumptions the probability matrix for the flavour oscillation of a cosmological source is as follows

$$\begin{pmatrix} \nu_e^{Earth} \\ \nu_\mu^{Earth} \\ \nu_\tau^{Earth} \end{pmatrix} = \frac{1}{18} \begin{pmatrix} 10 & 4 & 4 \\ 4 & 7 & 7 \\ 4 & 7 & 7 \end{pmatrix} \begin{pmatrix} \nu_e^{source} \\ \nu_\mu^{source} \\ \nu_\tau^{source} \end{pmatrix} \tag{1.28}$$

As the flavour ratio at the astrophysical source for proton-proton interactions is $(\nu_\mu : \nu_e : \nu_\tau) = (2 : 1 : 0)$ the expected flavour ratio at Earth due to neutrino oscillations is $(\nu_\mu : \nu_e : \nu_\tau) = (1 : 1 : 1)$

There are still some unknowns in the mass eigenstate hierarchy. As Δm_{ij}^2 is the difference in masses measured, there is no clear answer to the order of the masses. There are two possible models; in the normal hierarchy m_3 is the largest, in the inverted hierarchy m_3 is the smallest [10]. The best estimates of the parameters are:

$$\begin{aligned} |\Delta m_{31}^2| &= 2.40 \times 10^{-3}\,\mathrm{eV}^2 \sim |\Delta m_{23}^2| \\ \Delta m_{21}^2 &= 7.65 \times 10^{-5}\,\mathrm{eV}^2 \end{aligned} \tag{1.29}$$

The flavour ratio of neutrinos at the source will be different than that observed at Earth due to neutrino oscillations. Here we have described one method of oscillation due to the mixing of mass and flavour eigenstates resulting in an expected ratio at Earth of $(1 : 1 : 1)$ i.e. there should be equal numbers of muon, tau and electron neutrinos, if the source production mechanism is proton-proton interactions.

1.5.3 Neutrino Interaction

There is a small probability that neutrinos will interact with a nearby nucleon via the weak force. At high energies these are dominated by deep inelastic scattering in charged current and neutral current interactions. For neutral current interactions a neutrino will interact with a nucleon exchanging a Z^0 boson producing a neutrino and a hadronic shower

$$\nu + N \rightarrow \nu + X \tag{1.30}$$

where N is the nucleon and X is the hadronic shower component [10]. Charged current interactions involve a neutrino interacting with a nucleon exchanging a $W^\pm$ boson which then produces a charged lepton and a hadronic cascade. As stated above, neutrinos come in three flavours - electron, muon and tau - and the charged lepton produced in charged current reactions depends on the flavour of the neutrino as follows [8]

$$\nu_e + N \to e + X \tag{1.31}$$

$$\nu_\tau + N \to \tau + X \tag{1.32}$$

$$\nu_\mu + N \to \mu + X \tag{1.33}$$

The interaction cross section, as shown in Figure 1.3, is linearly proportional to neutrino energy up to about 10^4 GeV. Above this energy there is a change in slope due to the mass of the Z boson. At energies greater than 10^6 GeV the neutrino and antineutrino cross sections become equal because the contribution of valence quarks in the quark distribution functions is negligible. For energies above 10^6 GeV the Earth becomes opaque for neutrinos and thus no neutrinos will travel all the way through the centre of the Earth without interaction.

The Glashow resonance is the resonant formation of a W boson, and this occurs when anti-neutrinos interact not with nuclei but with electrons,

$$\bar{\nu}_e + e^- \to W^- \tag{1.34}$$

this reaction is resonant for a neutrino energy of 6.3 PeV. A signal at such an energy could be used to study the electron neutrino component at the source.

The low probability of low energy neutrino interactions means that they are difficult to detect as the detections of neutrinos is done via secondary particles. At higher energies the neutrinos are more likely to interact as they have a higher cross-section however that also means that they are more likely to interact before reaching the detector. In order to use neutrinos as astrophysical tracer particles a large detector volume is required to allow for enough neutrino interactions, see Chapter 2.

1.5.4 Atmospheric Neutrino Production

When cosmic rays interact in the Earth's atmosphere they produce a cascade of secondary particles in a shower. If the initial cosmic ray is a hadron the interaction will produce

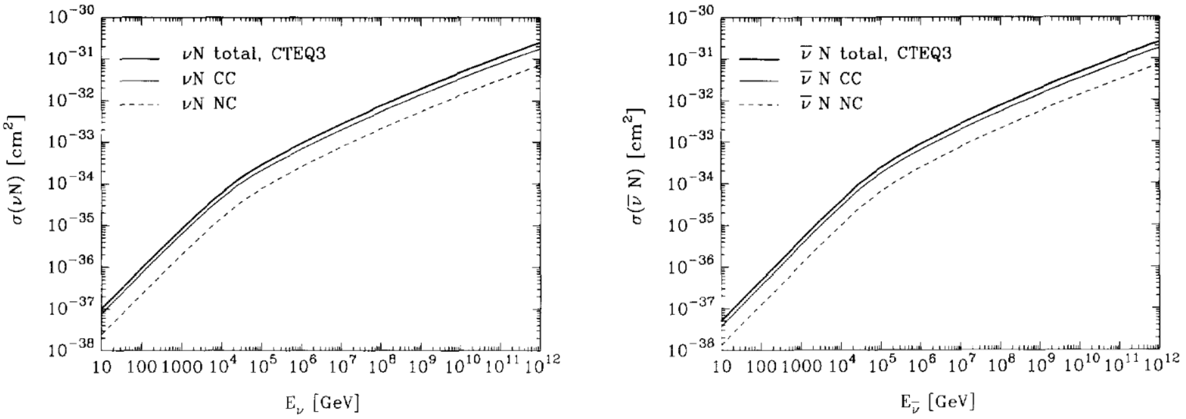


Figure 1.3: Cross sections for neutrino nucleon interactions, left for neutrinos, right for antineutrinos. Dashed lines are for neutral current interactions, thin solid lines are for charged current interactions, and the thick solid line is the total. Both interactions cross sections increase with energy allowing for detection of high energy neutrinos as they are more likely to interact. However that also means that they may not reach the detector before interacting. Figure from [19].

secondary mesons such as kaons and charged pions, which then decay to neutrinos. This type of atmospheric neutrino is referred to as 'conventional' and is the most abundant atmospheric background in neutrino experiments. At high energies neutrinos are only produced if the kaons and pions decay, not if the pions and kaons interact in the atmosphere. The ratio of decay to interaction probability for pions and kaons is approximately proportional to 1/E [20]. Due to this there will be an increase in interactions at high energies, thus fewer neutrinos will be produced and the energy spectrum for conventional neutrinos will be steeper than the cosmic ray spectrum, following $E^{-3.7}$ [10].

Another source of atmospheric neutrinos is from the decay of mesons containing charm and bottom quarks. These decays happen quickly before any further interaction and are known as 'prompt'. As the particles rarely have other interactions and will quickly decay the prompt flux will have the same energy spectrum as the cosmic ray energy spectrum at $E^{-2.7}$ [16]. There has been no clear detection of a prompt neutrino flux from our atmosphere but upper limits have been calculated [21].

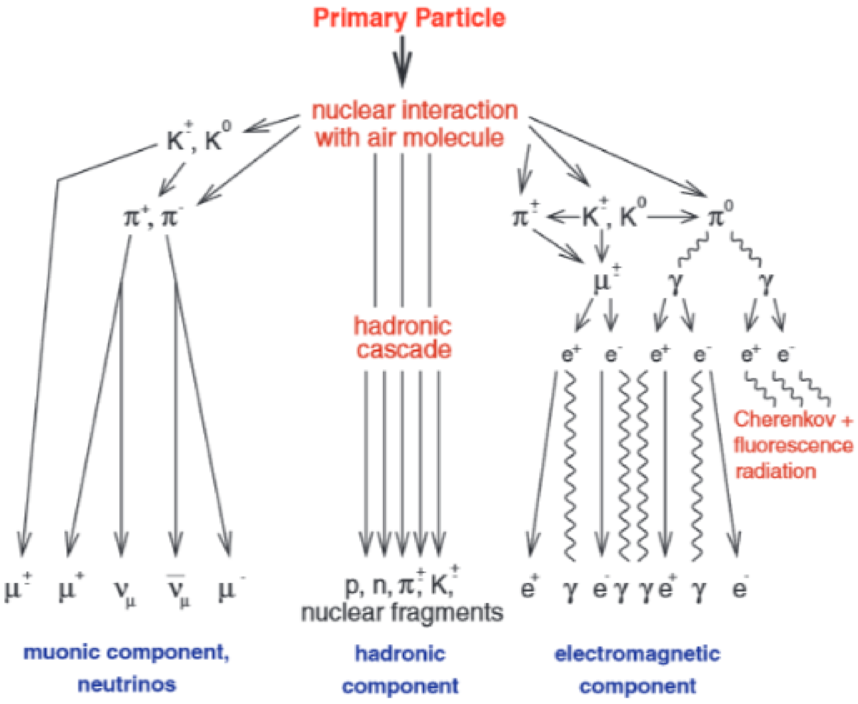


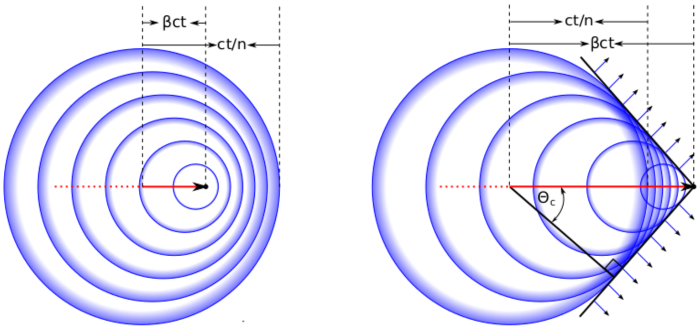
Figure 1.4: Cosmic ray interactions in the atmosphere causing a shower of secondary particles. There are three components to the shower, muonic component, hadronic component and electromagnetic component. The atmospheric neutrinos come from the muonic component producing muon neutrinos from pion and kaon decays.

1.5.5 Cherenkov Radiation

The detection of neutrinos is achieved by the detection of secondary particles created in neutrino interactions as described in the previous section. When a charged particle travels through a dielectric medium at a velocity larger than the phase velocity of light for that medium i.e. $v > c/n$ where n is the refractive index, then Cherenkov radiation is emitted at a characteristic angle

$$\cos\theta_c = \frac{1}{\beta n} \tag{1.35}$$

where $\beta = v/c$ is the Cherenkov threshold where once the phase velocity is greater than c will result in Cherenkov radiation [10]. Figure 1.5 shows a sketch of this process, with the characteristic angle shown.



(a) $v < c$ **(b)** $v > c$

Figure 1.5: Diagram of Cherenkov light production, where in (a) the velocity is below the threshold, and in (b) the velocity is greater than the speed of light for the medium.

The method for Cherenkov production is by the travelling charged particle polarising atoms and molecules along its path. As the atoms and molecules return to the ground state they emit dipole radiation; these photons can interfere constructively or deconstructively depending on the particle's velocity. According to Huygens' principle, light is emitted as spherical light shells; if the velocity is less than the phase velocity the shells interfere deconstructively. If the particle is faster than the phase velocity, constructive interference results in the emission of light at an angle described by Eq 1.35.

The spectral distribution of Cherenkov light per path length per wavelength interval is given by

$$\frac{dN}{dXd\lambda} = \frac{2\pi z^2\alpha}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2}\right) \tag{1.36}$$

where z is the charge of the particle in units of the elementary charge, λ is the photon's wavelength and α is the fine structure constant.

The refractive index for ice is $n \sim 1.33$; for relativistic particles emitting Cherenkov radiation in the ice the angle will be $\theta = 41$ degrees. The energy loss by Cherenkov light emission is $\sim 100\,\mathrm{eVcm^{-1}}$ and so is negligible compared to the total energy loss rate of the particle.

1.6 Summary

In this chapter we first describe high energy cosmic rays that can be detected at Earth that are produced and accelerated by astrophysical sources. We show the results for the energy spectrum of cosmic rays from multiple experiments and report the spectral index that has been measured between particular features of the spectrum which are believed to be caused by transitions between dominant astrophysical sources. We also describe the neutrino and gamma ray production that is believed to happen after cosmic rays interact with photons and gas. This allows neutrinos and gamma rays to be used as astrophysical tracers. The limitations for high energy cosmic rays are discussed, with ultra high energies needed for direction preservation and the GZK limit causing a cut off in the energy spectrum.

We introduce the neutrino as a tracer particle for cosmic ray acceleration sources and discuss the neutrino oscillation from astrophysical sources, with the expected neutrino flux and flavour ratio expected to be detected at Earth. We also discuss neutrino production in the atmosphere due to cosmic ray air showers with the expected spectral index for the flux. Neutrino interactions and their secondary particles are described with their interaction cross sections shown. Finally Cherenkov radiation is described as a means of detecting secondary interaction particles as they travel faster than the speed of light for the medium.

Chapter 2

The IceCube Neutrino Observatory

Using neutrinos as astrophysical tracer particles is challenging as neutrinos are hard to detect, and when a detection is made distinctions must be made as to whether it was of astrophysical or atmospheric origin. IceCube was the first operational large scale neutrino observatory and reported the first detection of an astrophysical neutrino spectrum. In this Chapter we describe its design and operation in addition to reporting current astrophysical results.

2.1 Early Neutrino Observatories

The first attempt at a large scale neutrino detector was the DUMAND experiment [22], which was planned to be a series of strings of photomultiplier tubes off the coast of Hawaii. During the first deployment in 1993 the pressure housing on the first string developed leaks and the string failed. In the same year the Baikal neutrino telescope [23] in Lake Baikal, Russia, was the first detector to have three strings and thus be able to do a full spatial reconstruction of a muon trajectory.

The first generation of neutrino detector to be located at the South Pole was AMANDA (Antarctic Muon And Neutrino Detection Array) [24] which utilised the thick clear ice under the South Pole. Deployment began in 1995 and was completed in 2000. The detector comprised 19 stings placed in the ice with a total of 667 optical modules with depths between 1500 m and 2000 m. During its development it was found that the strings were best deployed to deeper depths as above 1300 m bubbles in the ice cause a small scattering length making track reconstruction difficult [25].

2.2 Detector

The IceCube Neutrino Observatory is a neutrino detector located at the South Pole. Construction of the full detector was completed in December 2010, but partial configurations runs were done in previous years. Figure 2.1 shows the final detector configuration. The detector works by utilising the fact that neutrino interactions produce secondary leptons that will be travelling faster than the speed of light in the medium (in this case ice) and thus will emit Cherenkov photons, see Section 1.5.5.

A large detector volume is needed due to the small interaction cross section of neutrinos at low energies, it is also needed for higher energy neutrinos where cross-sections are higher but fluxes are lower. The ice cap at the South pole is 3km deep and thus provides a large quantity of interaction material with excellent optical properties. In addition, being located at the Amundsen-Scott South Pole Station allows for logistical support for construction and operation [26].

2.2.1 IceCube In-Ice Array

The in-ice array consists of 86 strings, of which 8 are DeepCore strings described in the next section, there are 60 Digital Optical Modules (DOMs, see below) on each string, placed at depths between 1450 m and 2540 m within the ice. The DOMs placed on each string have a vertical separation of 17m. The strings are in a hexagonal array with spacing of 125m between strings. The whole in-ice array encloses a cubic kilometre volume.

The DOMs record the Cherenkov photons from the secondary particles. There are two different signature event topologies seen from secondary particles: tracks and cascades. Of the secondary leptons produced the electron will not have a long track; instead an electromagnetic shower or cascade will be emitted outward from the interaction point (see Figure 2.2a). Because of this structure it will be harder to gain any directional information.

The track-like signature in Figure 2.2b comes from charged current interactions of

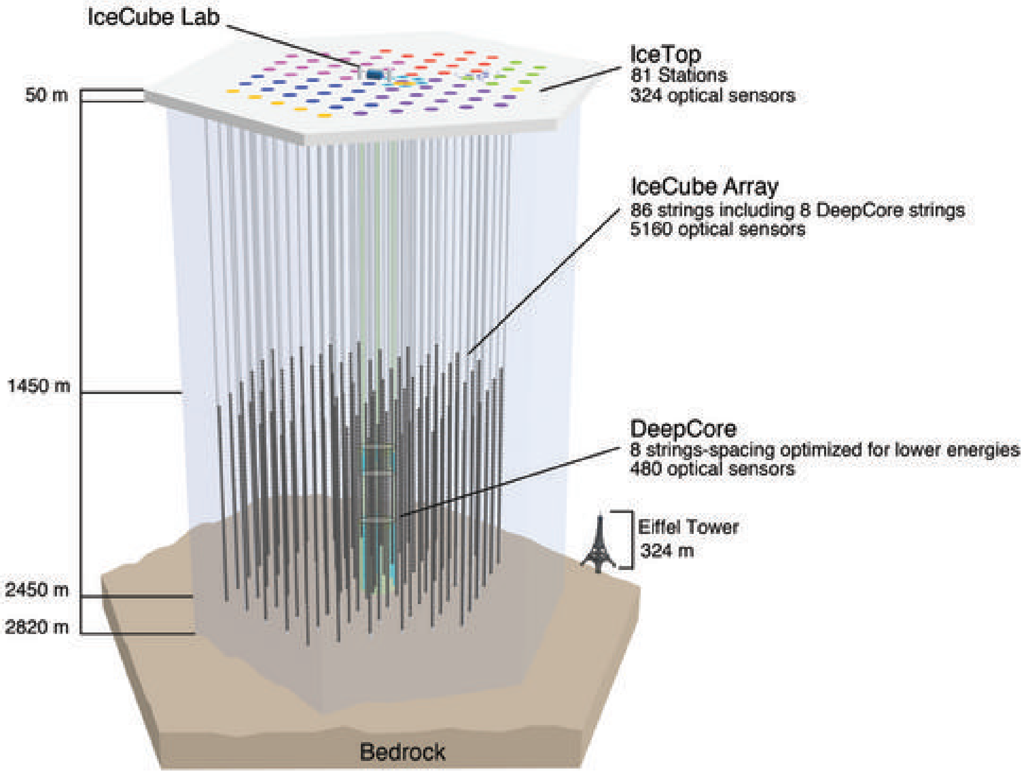
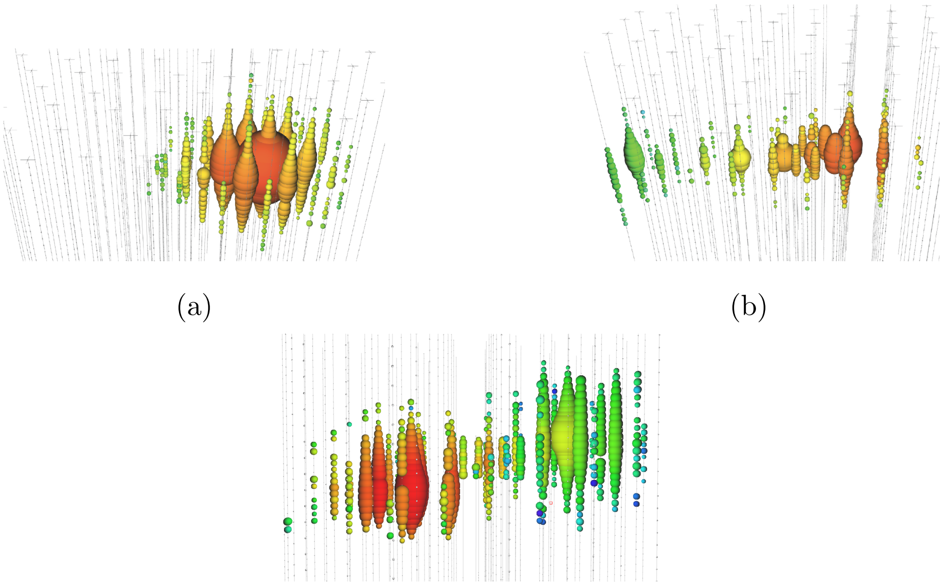


Figure 2.1: The IceCube Neutrino Observatory with in-ice array, DeepCore and IceTop.



(a) (b)

(c)

Figure 2.2: Three types of events as would be seen by the IceCube In-Ice array for an interaction of (a) an electron neutrino (b) a muon neutrino (c) a tau neutrino. Each point represents a DOM that has recorded photoelectrons with the radius of the point proportional to the amount of photoelectrons and the colour, from red to blue, indicating the timing of the detection. The electron neutrino is a single cascade with a central region and no outgoing particle to improve track direction. The muon neutrino has a clear track through the detector. The tau event has an initial cascade then a secondary cascade.

a muon neutrino, with an initial hadronic shower and a muon which travels outward preserving the direction of the original neutrino. This means tracks can have a high angular resolution for their incident neutrino direction. Due to this, secondary muons have a range of several kilometres with a clear track direction in the detector that can be reconstructed with a good angular resolution for its direction, therefore muons are the favorable lepton to use in analysis that require direction.

The final type of event is a tau neutrino interaction, shown in Figure 2.2c. Taus above $\sim 1\,\mathrm{PeV}$ will have a “double bang event” track with the starting interaction producing a cascade and a tau followed by a short range track then a tau decay cascade. Currently, no tau events have been identified individually as being detected by IceCube.

2.2.2 DeepCore

Within the detector is Deep Core, a denser array of eight strings at the centre of the detector. The string spacing varies from around 41 m to 104 m between each string. DeepCore has high quantum efficiency DOMs located below 2100 m where the ice is clearest.

Below 2100 m 50 high quantum efficiency DOMs are located on each string with a vertical spacing of 7 m. There is a gap where no DOMs were placed between 2000 m and 2100 m as there is a region of high dust concentration in the ice that increases scattering and absorption. Above 2000 m are a further 10 DOMs with 10 m vertical spacing that can act as a muon veto.

DeepCore has enhanced sensitivity to lower energy neutrinos below 10 GeV due to the dense spacing and increased efficiency of the DOMs [26]. DeepCore expands the range of energies IceCube is sensitive to and allow IceCube to detect atmospheric neutrino oscillations, as well as possible neutrinos from WIMP dark matter annihilation and Galactic supernovae.

2.2.3 IceTop

Co-located with IceCube is IceTop, a surface array of 163 Cherenkov detector tanks used to detect extensive air showers at the surface. The tanks are located on 81 stations situated above the IceCube strings, which are configured in a triangular grid pattern with individual spacing of 125 m, so that IceTop covers an area of 1 square kilometre, see Figure 2.3 [27]. The two tanks at each station are separated by 10 m. Each tank is filled with ice up to approximately 0.9m, and has two DOMs, one high gain and one at low gain to increase the range of air shower detection.

The IceTop array is able to detect cosmic ray air showers with primary energy in the PeV to EeV range, with the central denser tanks above DeepCore sensitive to 100 TeV [26]. IceTop is used for cosmic ray studies sensitive to the knee region of the energy spectrum. IceTop also acts as a partial surface veto for downward-going events in the in-ice array but this acts in a very small angle range.

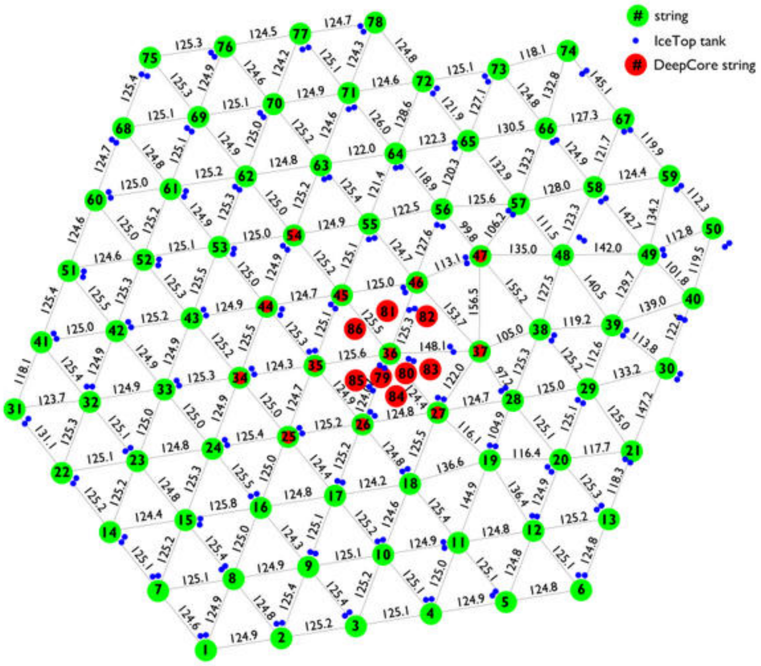


Figure 2.3: The IceCube Neutrino Observatory surface view with the in-ice array, DeepCore and IceTop.

2.2.4 IceCube Laboratory

The final installation located at the centre of the array is the central operations building- the IceCube Laboratory (ICL), see sketch in Figure 2.4. All surface cables from the in-ice and IceTop arrays are fed into the building through two cable towers and into a server room. The server room runs all the data acquisition systems as well as the at South Pole data filtering.

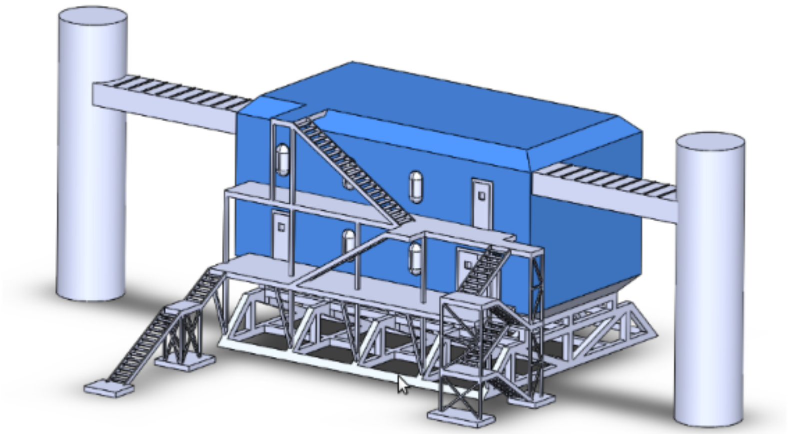


Figure 2.4: The IceCube Laboratory located at the South Pole. The two cable towers on either side serve to feed in the cables from the array into the server room.

2.3 Digital Optical Modules

There are a total of 5160 Digital Optical Modules (DOMs) deployed in the in-ice array. The DOMs consist of a 35.6 cm diameter spherical glass pressure resistant housing that contain a downward facing 25 cm diameter photo-multiplier tube (PMT) in the lower half and electronics in the upper half. The circuit boards for the DOM include the Main Board which controls data acquisition, communication, calibration and low-voltage power conversion. Other circuit boards generate PMT high voltage, interface to the PMT pins, delay PMT signals, and generate calibration light flashes that can be recorded by other DOMs.

The DOMs are designed to withstand the harsh conditions of temperature and pressure. DOMs operate in these conditions with precise recording of PMT pulse widths and amplitudes with nanosecond resolution. The DOMs were designed with long term reliability in mind and as of 2016 only 87 of 5160 deployed DOMs have failed. Of the failed DOMs, 55 died during deployment and thus were ruled to be due to cable damage, water leaks and freeze-in damage. The remaining 32 failed DOMs failed after deployment with no particular pattern of failure evident.

The PMTs can detect light from secondary interaction particles in the ice that could be up to 500 m away, in the energy range of 10 GeV to 10 PeV. Because of this range, the PMT waveforms can have an amplitude from 1mV up to the limit of the PMT which is approximately 2 V, with widths from 12 ns up to around 1500 ns. To provide for the range of signals each DOM includes multiple digitisers with overlapping dynamic range and different sampling speeds. Each DOM can detect a hit, which is where it records the individual photons using a PMT waveform that can include photons arriving up to 6.4 μs later. The time of the hit is saved along with the waveform shape, allowing for the determination of the times of arriving photons relative to this reference. The accumulated hit data is then transmitted to the ICL [26].

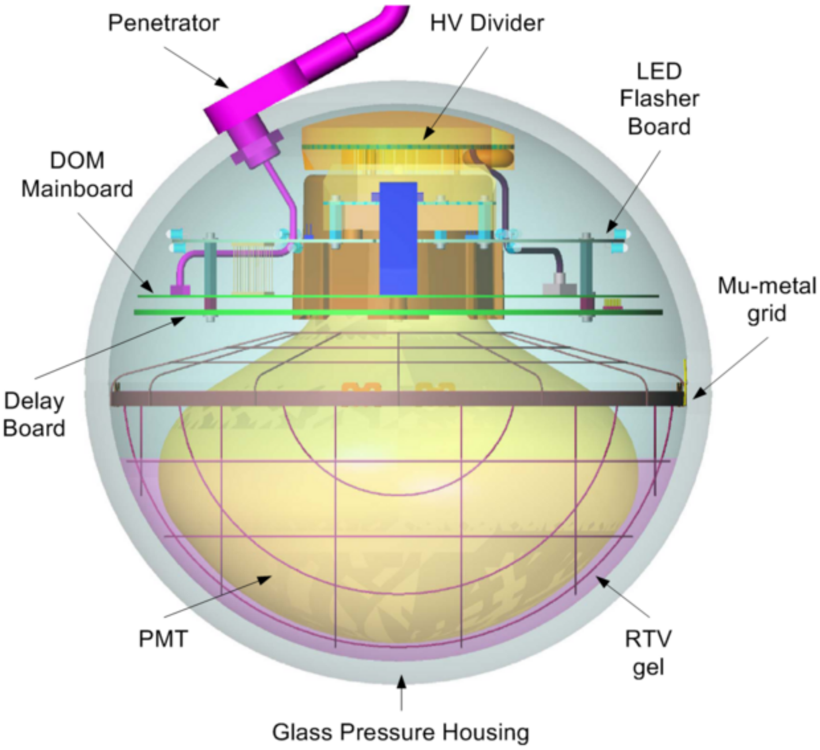


Figure 2.5: Schematic of a DOM, with the PMT pointed downwards and electronics at the top of the DOM.

2.4 Ice Properties

The secondary particles from the neutrinos cause Cherenkov photons that can be detected (See Section 1.5.5). The intensity of the photons is recorded by each DOM in the IceCube detector, but to understand where the light the DOM detects comes from, the nature of the ice around the DOMs needs to be understood.

Studies of the ice properties were conducted using both the AMANDA and IceCube detectors. Mounted within each DOM is an LED flasher board which is used to generate light in the detector that can be used for detector calibration purposes as well as ice properties studies. The LEDs emit light at $405 \pm 5\,\mathrm{nm}$ [26].

Photons can be scattered and absorbed within the ice; the amount by which this happens is important, as it affects how the IceCube detector operates. In deep ice scattering occurs on microscopic scattering centres, such as submillimetre scale air bubbles or micron scale dust grains. Any material in the ice with a different refractive index can cause scattering. The scattering in-ice can be calculated using Mie scattering calculations from Maxwell's equations. In the IceCube detector there will be multiple scattering that occurs, thus the scattering length measured is not the geometric scattering length, which would be the average distance between scatters. Instead it is represented using the effective scattering length, λ_e, which describes the length at which for a beam of light the cluster of photons moves along the incident direction in the medium with forward peak scattering occurring i.e. the photons are scattered but they are scattered in the forward direction of the incident photon. The length for this pattern is the effective length λ_e, this takes into account the occurrence of multiple scattering [25].

The absorption length λ_a is the distance at which the photon survival probability drops to 1/e. Molecular absorption in the pure ice can occur, however it occurs in the red and infrared, which is outside the wavelength range of the IceCube PMTs. For the wavelengths 200 to 500 nm pure ice is highly transparent and as such absorption is dominated by the dust in the ice.

Shown in Figure 2.6 are the results of a study using the AMANDA detector to measure the absorption and scattering within the ice. The study was done for wavelengths 313 to 560 nm at depths from 1100 to 2350 m using both continuous light and pulsed light sources.

At low depths there are air bubbles present in the ice but as the pressure increases with depth the air bubbles are compressed, resulting in a transition from gas phase to solid non scattering air-hydrate clathrate phase, and by about 1500 m all air bubbles have transformed into this state. The air bubbles were measured to result in very short scattering lengths for light, but at the transition of 1500 m they do not contribute as much of an effect to the scattering of light [25]. Hence at the depths of the IceCube detector the optical properties depend more on the concentration of dust in the ice.

The peaks in the scattering and absorption coefficients in Figure 2.6 at depths of greater than 1400 m peaks are dust concentration due to periods of increasing cold during the last glacial period in the late Pleistocene.

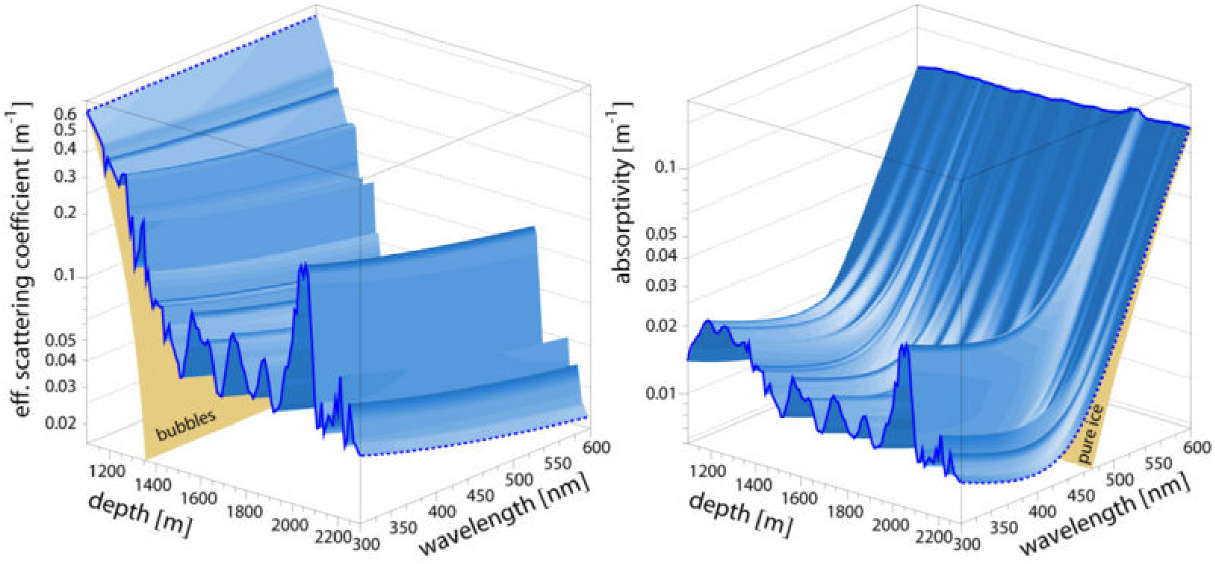


Figure 2.6: (*left*) The optical scattering (λ_e) as a function of depth and wavelength (*right*) The optical properties of absorption (λ_a) as a function of depth and wavelength.

Between 2000 to 2100 m there is a large increase in the amount of absorption and scattering because of a stratum of dust that is denser than any other. This stratum, known as the *dust layer*, can cause issues with the IceCube detector and special care is taken in analyses to minimise its impact.

There are additional effects on the optical properties of the ice, caused by the deployment of the strings, where the holes made in the ice using the hot water drill would refreeze around the strings and as the ice was new it had not been subject to the time and pressure to cause air bubbles to transition to air hydrates. As with the shallow depth ice this means that the air bubbles would affect scattering and absorption in the hole.

The information provided by the on-board flashers is used to make tables of the ice properties that can be used when reconstructing the event's direction and energy. By accurately taking into account scattering and absorption the error on the event reconstruction can be minimised.

2.5 Astrophysical Results

Since the start of its deployment IceCube has operated in partial, and since early 2011, in full configuration and has detected the first astrophysical neutrino flux.

The first confirmed detection of astrophysical neutrinos by IceCube was by using a High Energy Starting Events (HESE) analysis which used the edges of the detector volume to create a veto region to tag incoming muons [28]. Using this veto, events that had their neutrino interaction occur within the detector could be identified. This analysis was sensitive to the southern sky as down-going neutrinos, if atmospheric in origin, would have accompanying muons from the cosmic ray air shower, and thus would not pass the veto, as the cosmic ray produced muons would be seen entering the detector. What remain is a selection of starting events with high probabilities to be astrophysical. The skymap in Figure 2.7 shows the latest HESE results for 6 years of IceCube data. So far no significant neutrino point source nor correlation with known sources has been found.

The neutrino energy spectrum is an important astrophysical measurement to make and analyses have been done using increasing years of IceCube data [9], [30] and [21]. All of these searches used samples of up-going muon neutrino events, that is detections of muons from a zenith range $90^o - 180^o$ [21]. Such events would come from neutrino interactions within the Earth and would consist of three types of neutrino events; prompt and conventional from the atmosphere, and astrophysical from distant sources. Prompt and conventional neutrinos come from cosmic ray interactions with the atmosphere as described in Section 1.5.4. The expected energy spectrum of conventional neutrinos would be a soft spectrum of $E^{-3.7}$, the prompt spectrum, although it has not yet been observed, is expected to follow the cosmic ray spectrum of $E^{-2.7}$. Finally the astrophysical energy at 2018-05-06 17-16-15.png

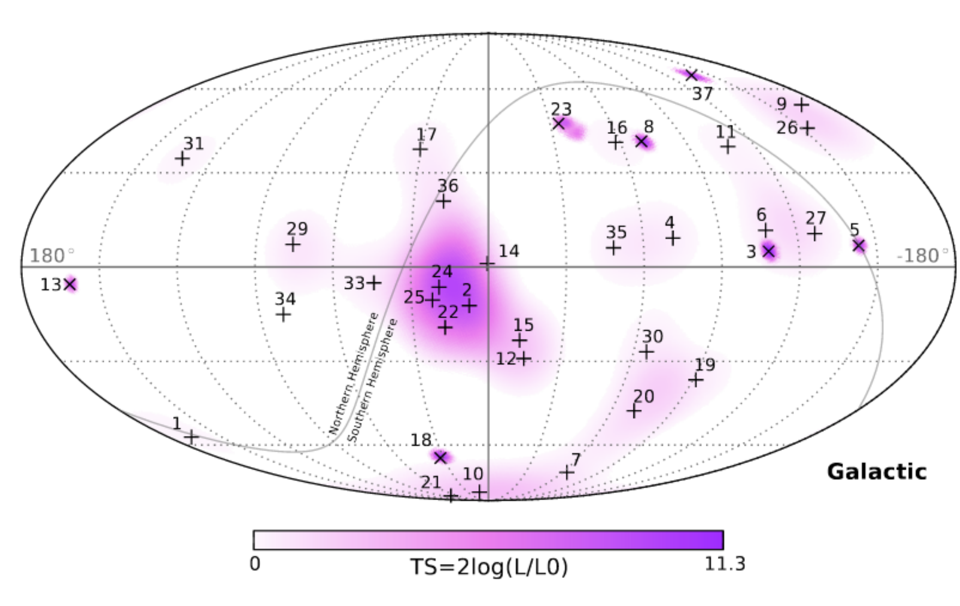


Figure 2.7: High energy starting event 3 year sample event direction. Crosses (+) are cascade events and x's are muon events. Shown is the test statistic of the point source clustering [29].

spectrum is expected to follow the cosmic ray spectrum at the acceleration source, E^{-2}. Thus the three components are expected to resemble that in Figure 2.8 with the detection of an astrophysical neutrino flux observable as an excess of events at high energies.

The most recent analysis of the neutrino energy spectrum in IceCube was done in [21] using a sample of 6 years of up-going muon events, with selection cuts made to only include high quality, well reconstructed events. This diffuse analysis involves fitting the data to the expected neutrino fluxes in bins of energy and cosine of zenith angle. Each bin is analysed by a maximum likelihood method where the expectation in each bin is a function of the physics and nuisance parameters. Physics parameters are variables the likelihood is trying to maximise the fit to, while nuisance parameters are any other variable in the fit needed to be accounted for but are fitted to a constrained range of values. In the analysis in [21] the likelihood per bin i is given by

$$\mathcal{L}_i = \left(\frac{\mu_i}{s_i/n_s}\right)^{s_i} \left(\frac{\mu_i}{d_i}\right)^{d_i} \tag{2.1}$$

where n_s is the ratio of live time for simulation and experimental data, s_i is the number of simulated events, d_i is the number of events in the data and μ_i is the expectation for

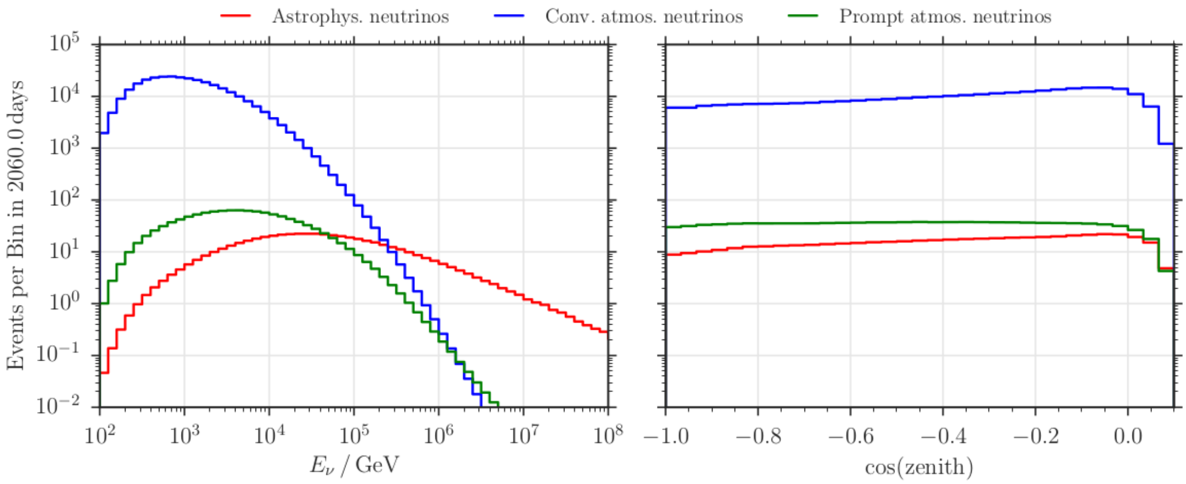


Figure 2.8: The distribution of the expected neutrino energy (left) and zenith angle (right) for the 6 year up-going sample in [21] based on the flux measurements in [31] and [32]. Shown are the expected astrophysical, conventional and prompt energy spectra. Astrophysical neutrinos will be detected as an excess of events at high neutrino energies.

the bin. The expectation μ_i depends on the signal θ and the nuisance parameters $\boldsymbol{\xi}$, and for this analysis it is the combination of the three neutrino fluxes (conventional, prompt and astrophysical) given by

$$\mu_i(\theta;\boldsymbol{\xi}) = \mu_i^{Conv}(\boldsymbol{\xi}) + \mu_i^{Prompt}(\Phi_{Prompt};\Delta\gamma_{CR},\lambda_{CR},\boldsymbol{\xi}_{det}) + \mu_i^{Astro}(\Phi_{Astro},\gamma_{Astro};\boldsymbol{\xi}_{det}) \tag{2.2}$$

where $\boldsymbol{\xi}_{det}$ includes all parameters taking into account neutrino detection uncertainties. The signal parameters are the astrophysical flux parameters Φ_{Astro}, γ_{Astro} and the prompt flux parameter Φ_{Prompt}. The nuisance parameters for this fit include the conventional flux,the Kaon-pion ration, cosmic ray spectral index $\Delta\gamma_{CR}$, the cosmic ray model λ_{CR}, the optical efficiency, scattering and absorption length and ice model.

The muon energy spectrum fit for different years of detector configurations are shown in Figure 2.9, with the fit to the data resulting in a combined flux with an astrophysical component clear at high energies. Note in the IC2012-2014 data there is an event with muon energy in the PeV range. This is the highest energy event observed by IceCube to date and is the subject of an energy reconstruction analysis in Chapter 7.

The final neutrino energy spectrum for all years is given in Figure 2.10 which shows the best fit for the conventional and astrophysical component, with the fit finding no

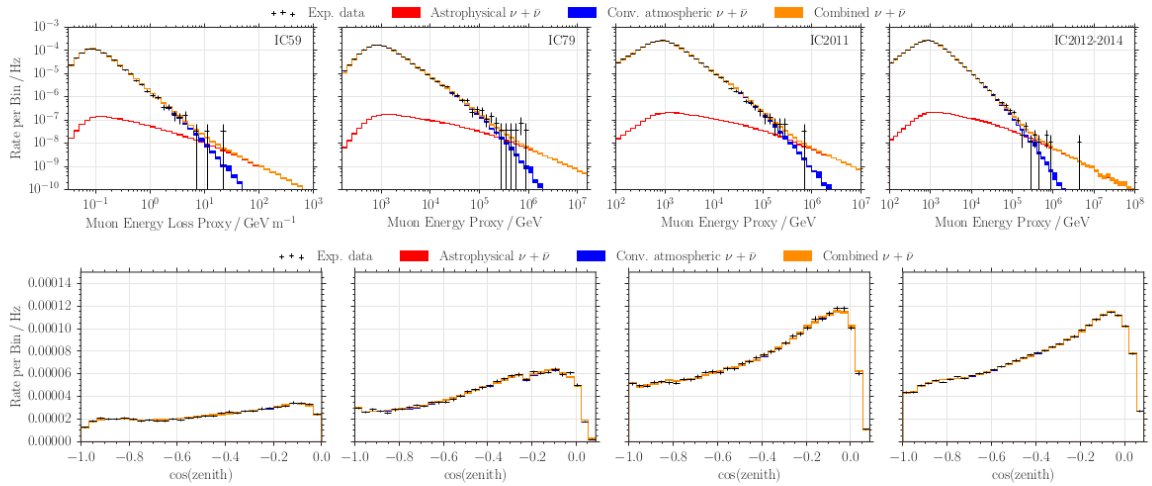


Figure 2.9: Distribution of experimental up going muon data for four separate detector configurations over six years for the muon energy proxy *(top)* and the reconstructed zenith *(bottom)* for the multi-year diffuse flux analysis in [21]. There is an excess in high energy events that results in a non-zero astrophysical component, indicating astrophysical neutrinos in the sample.

contribution from the prompt component for which upper limits are reported. Also shown is the starting events flux presented in [33]. The best fit values for the physics parameters are: spectral index of $\gamma = 2.13 \pm 0.13$, $\Phi_{Astro} = 0.9 \pm 0.28 \times 10^{18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$, $\Phi_{Prompt} = 0.0 \times 10^{18}\,\mathrm{GeV}^{-1}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{cm}^{-2}$ with upper limits $0 - 0.19$. The results of this analysis do not find a spectral index consistent with IceCube analyses using starting events, indicating there may be a spectral break or an additional astrophysical component at lower energies.

The reconstruction of the event's energy is an important component of a diffuse flux analysis as this analysis relies on the slight separation between the energy spectrum of conventional and astrophysical neutrinos. If the energy of muons and thus the neutrinos had a better resolution it would be easier to separate the different components of the spectrum.

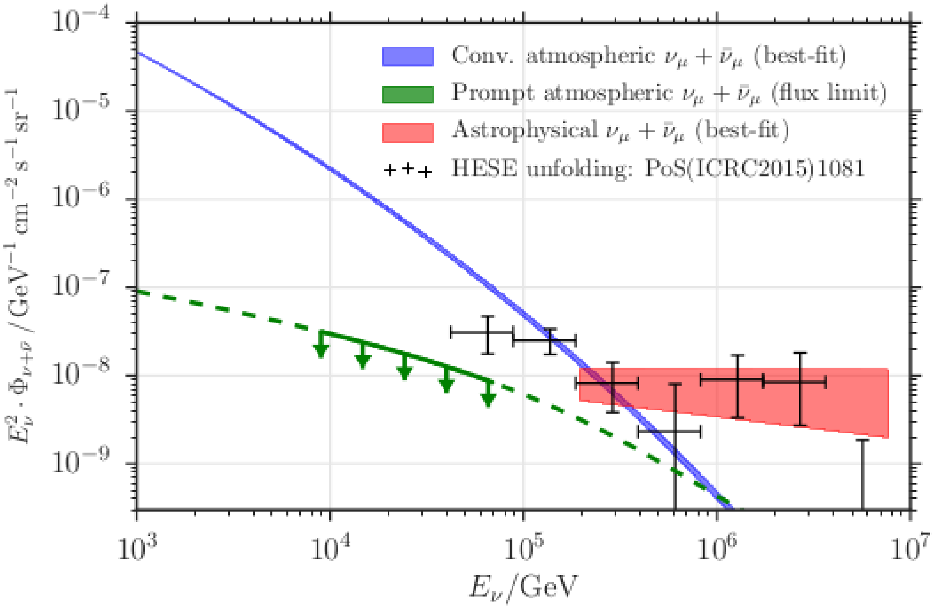


Figure 2.10: Best-fit neutrino spectra for the unbroken power-law model. With a conventional atmospheric component (*Blue*) and a astrophysical neutrino component (*Red*), with the width of the *blue* and *red* lines representing one sigma error in best fit values. The prompt component upper limit *Green* is shown. Black data points show the unfolding of the energy spectrum of the starting events analysis [33].

2.6 Summary

In this chapter we describe the IceCube neutrino observatory operation for detecting neutrinos from secondary interaction particles. The effect that the ice has on the detection medium for IceCube has been discussed, outlining the different properties at various depths of ice.

The first astrophysical neutrino detections have been achieved using IceCube in both a starting event analysis and an up-going astrophysical flux analysis. The results show the direction and energy of astrophysical neutrinos and can be used to perform neutrino astronomy for the first time. The neutrino flux gives important information on the nature of the acceleration source of cosmic rays, if neutrinos are produced near the source as is theorised. The diffuse flux analysis can be improved with a better resolution on reconstructed energy, as it relies on differentiating the three neutrino components by analysis of their energy spectra.

Chapter 3

Muons

Neutrino interactions of both astrophysical and atmospheric origin produce muons. Muons are also produced in cosmic ray air showers and these are the most commonly detected particles in the IceCube detector. Muons produce tracks within the detector that can be used for reconstruction of muon energy and for reconstruction of the muon direction, which are correlated to the neutrino direction for high energies. Muons are the favored lepton for use in a variety of analyses such as neutrino point source searches, diffuse extraterrestrial neutrinos, as well as standard and non-standard neutrino oscillations [34]. Of increasing importance in analysis is the use of starting muons in down-going neutrino event searches using a veto region within the detector to eliminate the atmospheric background. This chapter describes muon interactions and energy loss within the ice. Also described are the current muon energy reconstruction (Truncated Energy) and the event reconstruction (Millipede) that are used in IceCube.

3.1 Muon Energy Loss

Muon interactions in the detector will cause Cherenkov photons to be emitted by the charged particles resulting from the interactions, as well as from the muon itself. These interactions or energy loss mechanisms include ionization as well as radiative processes such as Bremsstrahlung, pair production and photonuclear reactions [8]. The expected energy loss rate per depth for a muon is approximately

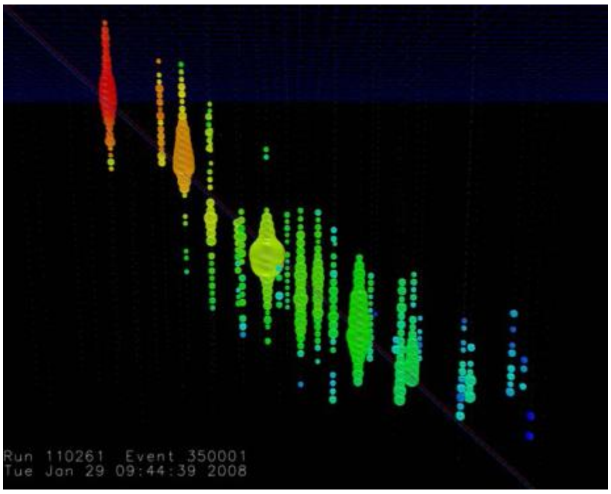


Figure 3.1: A muon event travelling through the IceCube detector. Each point represents a DOM that has recorded photoelectrons with the radius of the point proportional to the amount of photoelectrons and the colour from red to blue indicating the timing of the detection.

$$\frac{dE_\mu}{dX} = a + bE_\mu \tag{3.1}$$

where a represents the ionization loss which is approximately independent of muon energy, and b is the sum of the fractional energy loss of the radiative processes that cause stochastic energy losses i.e. $b = b_{Brem} + b_{pair} + b_{ph}$ [34]. The energy loss is dominated by ionisation at low energies, while at higher energy stochastic losses occur more often as shown in Figure 3.2. There is a critical energy where the ionisation and radiative losses contribute equally i.e. $E_c = a/b$ where $dE_\mu/dX = 2a$. Above this energy of approximately several hundred GeV the radiative losses are dominant.

The radiative processes cause stochastic energy losses proportional to the energy of the muon and are represented by the bE_μ term i.e. $b = b_{Brem} + b_{pair} + b_{ph}$. The contribution of different processes in Figure 3.3 shows the variation in their contribution with muon energy. These contributions also differ depending on the medium through which the muon travels.

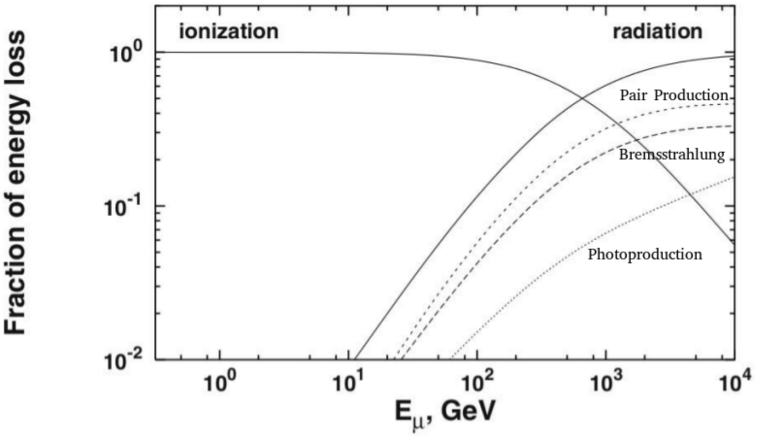


Figure 3.2: The relative importance of the different energy loss processes per g/cm^2 as a function of muon energy. The solid lines represent the cumulative distribution for ionisation and radiative processes. The short dashed curve is for pair production, the long dashed curve is for bremsstrahlung and the dotted curve is for photonuclear reactions. [8].

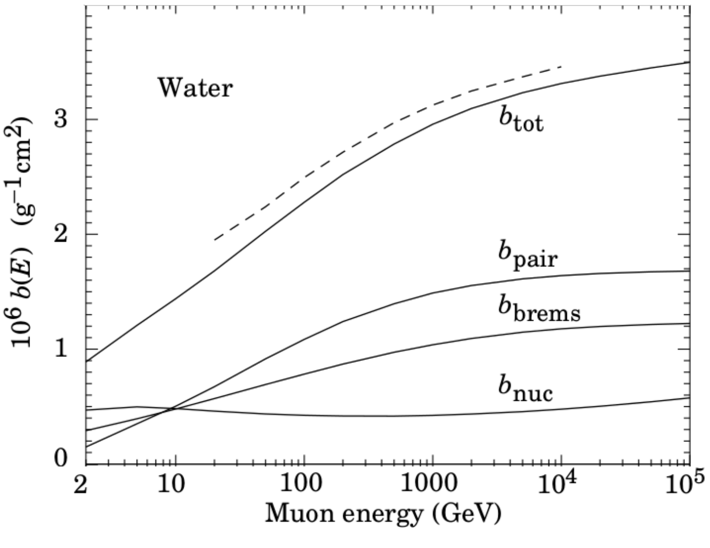


Figure 3.3: The energy loss contributions of muons in water from the radiative processes pair production, bremsstrahlung and photo nuclear reactions [35].

3.1.1 Bremsstrahlung

Bremsstrahlung is a stochastic process where the charged particle interacts with the electromagnetic field of a nucleus resulting in the particle being decelerated, and the subsequent emission of a photon, shown in Figure 3.4. In the case of muons, the emitted photons may cause an electromagnetic cascade in the detector. The energy loss for bremsstrahlung is proportional to the cross-section of interactions which is proportional to the energy of the muon.

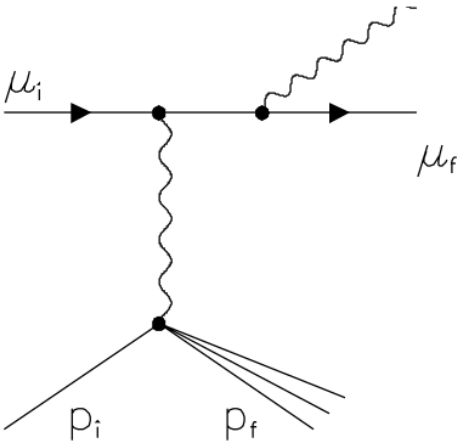


Figure 3.4: Feynman diagram of bremsstrahlung with the incoming muon μ_i interacting with the nucleus p causing the emission of a photon [36].

The cross-section for bremsstrahlung described in [36] is used for simulation (See Section 3.6) and is derived from [37] [38] [39].

3.1.2 Pair Production

Pair production, as shown in Figure 3.5, happens when the muon interacts with nucleons producing a virtual photon which then pair produces an electron and positron; e^+e^-. Pair production is a less stochastic process of energy loss than Bremsstrahlung and it therefore occurs more consistently along the muon track. The energy loss for a muon due to pair production is proportional to the energy of the muon. Thus higher energy muons will have a higher energy loss.

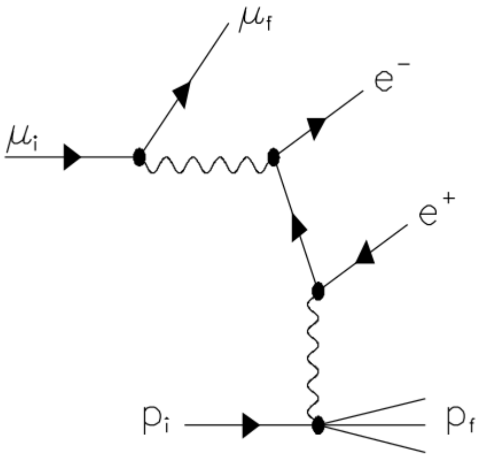


Figure 3.5: Feynman diagram of electron pair production due to interaction with a nucleus [36].

The cross section as first calculated in [40] was used in the simulation software outlined in [36] .

3.1.3 Photonuclear Reactions

Photonuclear reactions occur when a photon interacts hadronically with a nucleus and generates secondary hadrons, which produce hadronic cascades in the detector, shown in Figure 3.6. The contribution of photonuclear reactions to total energy loss increase linearly with energy and only have a substantial contribution at higher energies [8].

The cross section for photonuclear reactions is described in [41]; its implementation is outlined in [36]. The energy loss contribution is calculated from numerical integration of the differential cross section as described in [35] and [42].

3.1.4 Ionisation

A charged particle can also lose energy by excitation and ionisation of the atoms in the matter it is travelling through.

The ionisation energy loss is given by the Bethe-Bloch formula

$$\frac{dE}{dX} = \alpha^2 2\pi N \lambda_e^2 \frac{Z m_e}{A\beta^2} \left\{ \ln \frac{2 m_e \beta^2 \gamma^2 E'_m}{I^2(Z)} - 2\beta^2 + \frac{1}{4}\frac{E'^2_m}{E^2} - \delta \right\} \tag{3.2}$$

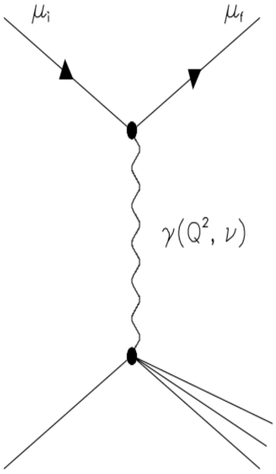


Figure 3.6: Feynman diagrams of photonuclear reactions with a nucleus [36].

where $\alpha = 1/137.036$ is the fine structure constant, N is Avogadro's number, Z is the atomic number, A is the mass number of the atoms in the medium, m_e and m_μ are the rest mass of an electron and muon, $\beta = p/E$ where p is momentum, $\gamma = E/m_\mu$, λ_e is the Compton wavelength of the electron, $I(Z)$ is the mean ionisation potential in the medium, $E_m^{'}$ is the maximum energy transferable by the electrons and δ is the density correction [42]. This process is also included in the simulation code in Section 3.6. For a muon propagating through ice the average ionisation energy loss is 200 MeV/m.

3.2 Muon Propagation

The average depth through the Earth that a muon of energy E_μ can penetrate can be calculated approximately via

$$R(E_\mu) = \frac{1}{b} \ln\left(\frac{bE_\mu}{a} + 1\right) \tag{3.3}$$

which is a variation on Eq 3.1, where a is the ionisation energy loss, and b is the fractional radiative energy loss per g/cm^2. This calculation works under the assumption that all of the muon energy losses are continuous. This assumption is true when ionisation is the dominant energy loss process, which is the case of the medium of rock in Figure 3.2 up to approximately 100 GeV. Above this energy the radiative processes are more dominant and cause fluctuations in the total energy loss creating a distribution of ranges. The maximum depth a muon can achieve is further than the average $R(E_\mu)$ if they do not undergo stochastic radiation losses. However as most muons do undergo stochastic losses, the depth they can reach is lower than $R(E_\mu)$.

The average of the range distribution is $\langle R_\mu \rangle$ for muons and is used to estimate the effects of fluctuations. The ratio of $\langle R_\mu \rangle / R$ is energy dependent with a higher energy having more radiative losses and thus a lower ratio [8].

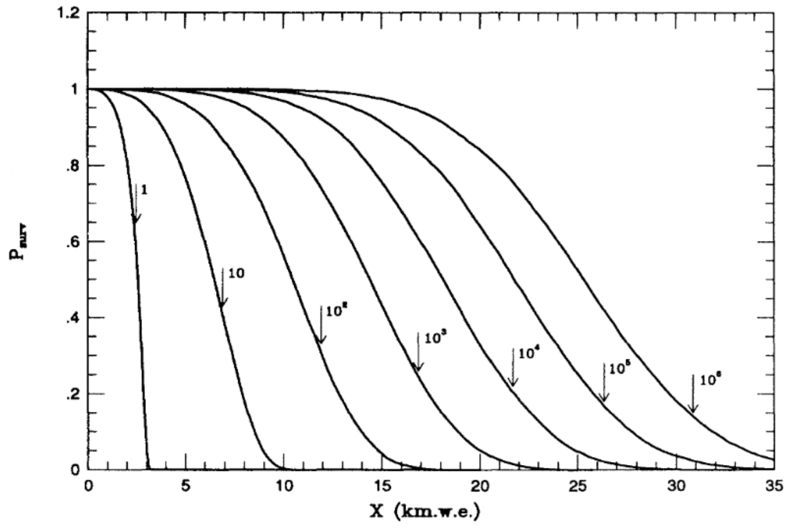


Figure 3.7: The survival probabilities of muons in standard rock. Each curve represents a different muon energy with the arrow indicating the depth of the average energy loss. Figure from [43]

3.3 Conventional Muon Energy Reconstruction

The energy of a muon propagating through the detector can be determined by analysis of the energy loss. IceCube finds the energy loss by using the Cherenkov light produced by the particle interactions (see Section 1.5.5) with the amount of Cherenkov light produced being directly proportional to the total energy present [44]. Figure 3.8 shows a simplified energy loss pattern as might be recorded by the IceCube detector, with a series of energy losses seen over the path length of the muon track. For muons below 300 GeV, if the muon track is contained within the detector, i.e. the muon track's starting and stopping points can be determined, then the energy is proportional to the path length. If a detected muon is contained within the detector the path length can be used to calculate the energy - for example, a 300 GeV muon has a probable mean path length of 1 km in-ice. For energies in the TeV range the specific energy loss dE/dX can be measured and used to determine the muon energy, E_μ. This can be done by using Eq 3.1 or by calculating the correlations between dE/dX and E_μ by Monte Carlo simulations [34].

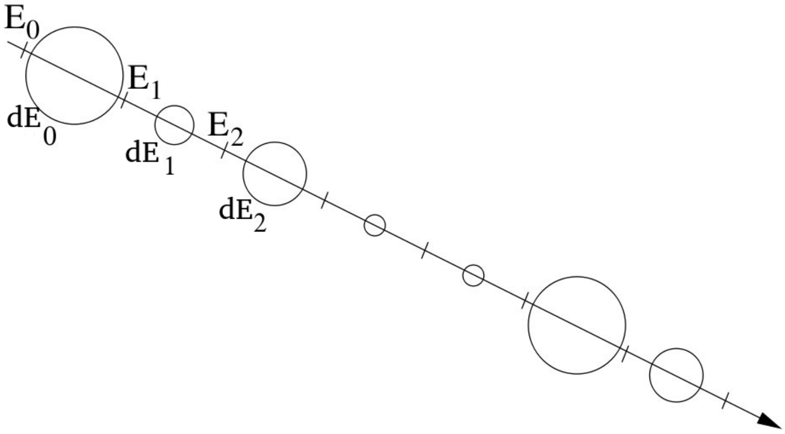


Figure 3.8: Visual of a muon energy loss pattern with initial energy E_0 and a series of energy losses dE_N. The radius of the circles is proportional to the energy loss in this display.

For muon energies above 1 TeV, muons lose most of their energy stochastically and this causes a larger spread in dE/dX values with a skewed mean.

The conventional dE/dX approximation for muons above 1 TeV is calculated by using the relationship that the yield of photoelectrons is expected to be directly proportional to the muon energy loss rate. This conventional dE/dX method is performed by summing the number of photoelectrons observed by the DOMs. Then an expected number of photoelectrons from the same track for a fixed energy loss of 1 GeV/m is calculated. By the earlier relationship for the energy loss, the calculated $dE/d\mathrm{x}$ values are equal to the ratio of the observed number of photoelectrons to the expected number for the 1 GeV/m track. Using this method a scatter plot of multiple simulated muon energies vs the calculated dE/dX is produced in [34], shown in Figure 3.9, which has a spread in dE/dX that is due to large stochastic losses. Fitting a curve to the scatter plot produces a relation to calculate E_μ from dE/dX [34].

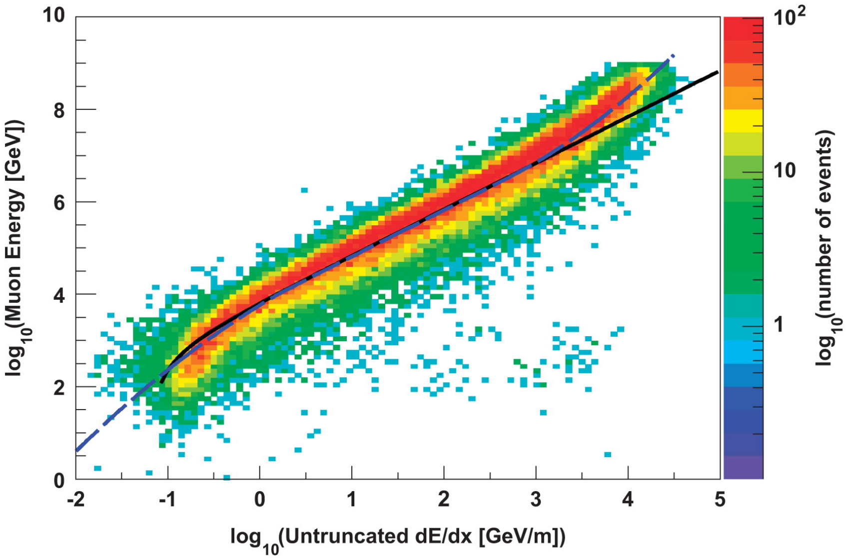


Figure 3.9: Simulated reconstruction distribution of true muon energy vs the conventional (untruncated) dE/dX.

3.4 Truncated Mean Energy

The truncated mean energy method (TE) is the primary muon energy reconstruction that is used in IceCube. It is based on the conventional method described in the previous section, but it divides the entire track into bins and dE/dX is calculated for each bin, as described above. Then a fraction of bins with the highest values of dE/dX are removed and a new average for dE/dX is calculated by summing the remaining observed photoelectrons and expected photoelectrons [34]. The large losses need to be removed to reduce the spread in dE/dX and improve the reconstruction resolution.

The TE has two different variations; the binned method and the DOMs method, which are explained in detail in the following sections.

3.4.1 Binned Method

The binned method of TE operates by taking the track of the muon event inside a detector such as IceCube then dividing the DOMs into 120m bins that are perpendicular to the track, shown in Figure 3.10. The bin size could vary depending on detector configuration along with the optical properties of the medium such as absorption and scattering of light. For IceCube, 120m was found to be the best binning in order for each bin to be independent of the others.

Each bin of the event has a calculated dE/dX from the ratio of observed photoelectrons to the expected photoelectrons for 1GeV/m energy loss within each perpendicular slice in Figure 3.10. Since DOMs that are further from the track have the increased likelihood of the sensor hits being noise, only DOMs within 0 m to 80 m of the muon track were used. DOMs were also removed from the bin if they did not have any hits, i.e. there were no recorded photoelectrons above threshold.

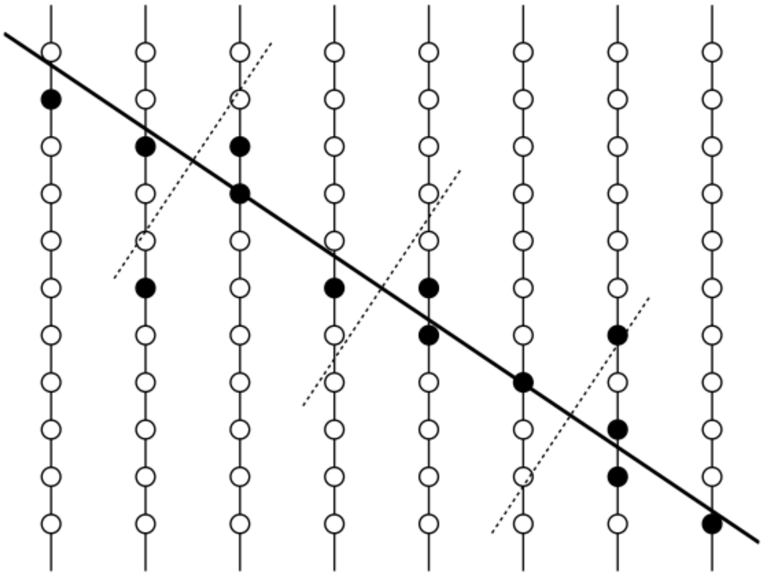


Figure 3.10: Diagram of the separation of detector DOMs into bins along the muon track, black dots represent active DOMs, dashed lines are the 120m bins [45].

The resolution for this method is summarised in Figure 3.11, for various cases, the method with the optimum resolution had 40% of the highest dE/dX bins removed. A minimum of three bins is required for the method to work as a 2 bin event would not have enough bins to remove at 40% truncation.

There is also another variation to this method where unhit DOMs were included in the bins. The method using the unhit DOMs is denoted as `AllBINs` because they used all DOM in the bins.

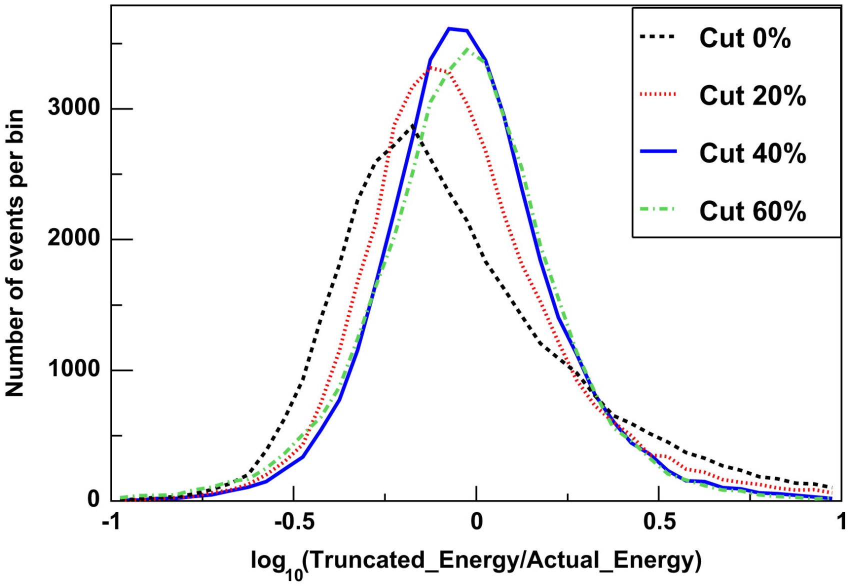


Figure 3.11: Truncated energy (TE) with various percentage cuts showing 40% with the least spread. Figure from [34].

3.4.2 DOMs Method

An alternative to binning the DOMs together is where each individual DOM can be used as its own bin. Each DOM in the detector has its own dE/dX calculated from dividing the observed number of photoelectrons by the expected number of photoelectrons for 1 GeV/m of energy loss. This method was also optimised to only allow a minimum of eight optical modules within 0 to 60m of the track. The truncation for this method removes the DOMS with the highest 50% of dE/dX. The remaining DOMs then have their dE/dX values averaged. The removal of 50% of DOMs was found from simulation to have the greatest effect on resolution.

The advantages of the DOMs method is that it allowed muon tracks to be shorter, as there does not need to be a minimum number of bins, and that additional data was used in the energy determination.

The DOMs method also has a variation to its method where DOMs that had recorded zero photoelectrons (no hit DOMs) were removed before the removal of the highest 50%. This was seen as advantageous to a wider range of energies. The method that included the unhit DOMs (`AllDOMs`), was reported as being the best for lower energy. However in later chapters of this work the comparisons to the TE method come from `AllDOMs`, as a detailed analysis of simulations found this to be the TE method with the best resolution, see Appendix B.

3.5 Muon Energy Loss Reconstruction: Millipede

A simple approach to reconstructing energy losses is spacial separation as outlined in [45] where the PMTs are separated into individual time bins and fit for the energy loss in the time bin to a template loss. This would not be accurate as fluctuations in muon energy loss as discussed in Section 3.1 occur over distances scales much smaller than the separation of IceCube strings and the light from a single energy loss can travel beyond a single PMT. Therefore the photons from one loss could be detected by DOMs in multiple time bins. Thus a segmentation method as discussed in the TE binning method (Section 3.4) could not provide a muon energy loss pattern reconstruction with resolution as would be needed for the body of this work.

The reconstruction method employed instead is Millipede - a segmented reconstruction method that utilises the fact that every PMT waveform is a linear combination of light emitted from all points along the track. This can be visualised in Figure 3.12 where each loss E_n has light detected by all PMTs. Because the stochastic energy loss of a muon track can take the form of electromagnetic showers, the Millipede algorithm can then use a linear unfolding which approximates each loss as different energy showers or cascades along the track [45].

There is a linear scaling between the light emission in a shower and its energy loss. As such cascades can be used as units of energy reconstruction by scaling the expected light output of a template cascade to that of observed data in each segment. The cascade reconstruction for the deposited energy in the shower is performed by comparing the observed number of photons in the PMT to the expectation of the light yield, Λ, for a template event.

For track events this principle can be extended along time. In IceCube the output of the PMT is collected using waveform-recording digitisers. For use in this algorithm the waveforms need to be unfolded into a number of photons per unit time $\vec{k}$. The relationship between the expected light and the templates can be given by:

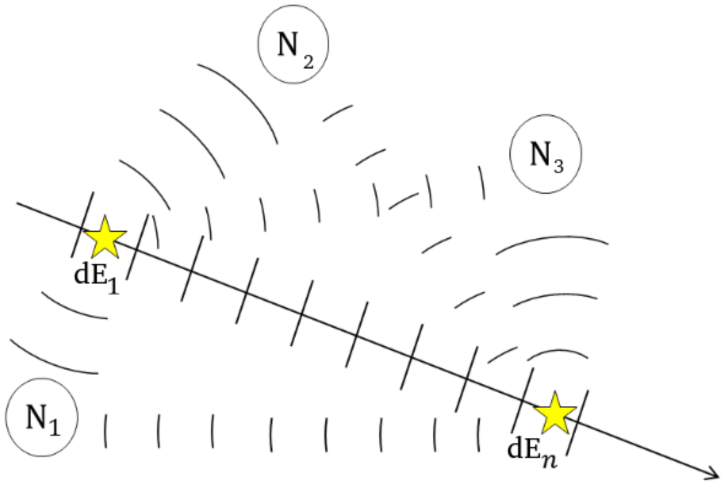


Figure 3.12: Muon energy loss along a track can propagate from loss dE_1 to DOMs N_1,N_2 and N_3.

$$\vec{k} = \mathbf{\Lambda} d\vec{E} + \vec{\rho} \tag{3.4}$$

where $d\vec{E}$ is the energy deposited along the muon track, $\vec{\rho}$ is the noise rate in the detector, and $\mathbf{\Lambda}$ is the predicted light yield per unit energy loss at every point in the detector from every source position at some reference energy. The parametrisation of the light yield in IceCube, $\mathbf{\Lambda}$, is done using a multi-dimensional spline surface [46] that has been fitted to multiple source Monte Carlo simulations.

Negative energies are unphysical and should not be allowed in the solution to the reconstruction. To ensure this, the algorithm uses a Non-Monotonic Maximum Likelihood (NMML) algorithm which achieves a resolution of approximately $10 - 15\%$ on the total deposited energy. The energy loss reconstruction of the Millipede algorithm is performed on a previously reconstructed track direction. The direction reconstruction used in this work is SplineMPE, which uses tables of the DOM response to energy for the possible track trajectory, which relies on the ice properties such as scattering and absorption. SplineMPE performs a maximum likelihood taking into account the geometry of the Cherenkov cone to reconstruct the track direction with good angular resolution.

3.6 Propagation software PROPOSAL

A simulation of the propagation of secondary particles through matter with an accurate description of the lepton energy losses is required for any neutrino experiment. For this work a simulation that has a stochastically correct description of muon interactions was needed. The Monte Carlo code PROPOSAL (PRopagator with Optimal Precision and Optimised Speed for All Leptons [36]) was used in this work. PROPOSAL is the successor to the Muon Monte Carlo (MMC) code which was the standard within IceCube for some time. PROPOSAL was written to provide precision on the magnitude and direction of energy losses along with their stochastic behaviours while also having increased speed based on numerical precision. Its language, C++, allowed it to be more easily integrated into the IceCube simulation chain [36].

PROPOSAL uses the same cross sectional formulae as MMC, with cross sections for ionization, Bremsstrahlung, pair production and photonuclear reactions as outlined in Section 3.1 with the cross section formulae given in [36]. All energy losses have continuous and stochastic components. In an ideal case all losses should be treated as stochastic, however that would require a very large amount of computations and below certain energies stochastic and continuous losses become indistinguishable. The artificial division between the stochastic and continuous losses is set by the input of an energy cut E_{cut} or a relative energy cut V_{cut}, where $V = dE/E$, here dE is the energy loss of the muon and E is the energy of the muon. The software is capable of calculating the propagation of muons from their rest mass 105.7 MeV to above 10^{11} GeV. PROPOSAL is capable of calculating the propagation of particles in three dimensions while taking into account Moliere scattering on atomic centres, which serves as an approximation to true muon scattering.

The results of PROPOSAL and MMC are regarded as identical and so the naming convention of the simulation was kept as MMC. In this work any simulation denoted as MMC came from a PROPOSAL simulation. Also note that in Chapter 5 an open source variation of PROPOSAL is used. In Chapters 6, 7 and 8 the IceCube version is used.

3.7 Summary

Muons are important in neutrino astrophysics as they are the most easily detected secondary interaction particle arising from neutrino interactions. These are produced in both the interaction of astrophysical neutrinos and atmospheric neutrinos as well as being produced in cosmic ray air showers. In this Chapter we discuss the properties of muon energy loss and propagation. A muon loses energy stochastically as it moves through the medium resulting in an uneven pattern of energy losses along its track. The energy loss pattern in the IceCube detector can be reconstructed via the Millipede method which uses a maximum likelihood unfolding of the waveforms in the DOMs and will be used for the energy loss patterns in Chapters 6, 7 and 8. The method used to simulate muon energy losses, PROPOSAL, is described with reference to its use in Chapters 5, 6, 7 and 8.

The current muon energy reconstruction method employed by IceCube is the truncated mean energy method (TE), which has two variations in its binning of dE/dX, but both remove the largest of the losses from the final set of losses used to calculate dE/dX. This produces good energy resolution for high energy muons but it removes $40 - 50\%$ of the energy losses in the pattern. The body of this work aims to create a muon energy reconstruction method that performs with better resolution than TE by using all the information in the energy loss pattern.

Chapter 4

Characterising Energy Reconstruction Resolution

In order to analyse and compare the performance of different energy reconstruction methods there must be a consistent comparison measure used. In this Chapter we describe the resolution measures developed and tested during this research and establish that the resolution measure allows for comparison of energy reconstruction performance as a whole. The resolution measures will be tested against idealised example reconstruction resolutions to demonstrate their limitations and effectiveness.

The measures described in this chapter to assess the reconstruction resolution are used in all later chapters to analyse the energy reconstruction performance of the truncated energy (TE) and Edepillim methods.

4.1 Measuring Reconstruction Resolution

A reconstruction resolution in an idealised energy reconstruction is likely to resemble that of Figure 4.1, where the observed result vs the true value is shown, which in terms of reconstruction is the reconstructed muon energy E_0 vs the true muon energy E_{TRUE}. In this example there is a slope of 45 deg between the true energy and the reconstructed energy, however this will not always be the case and such examples will be explored in later sections.

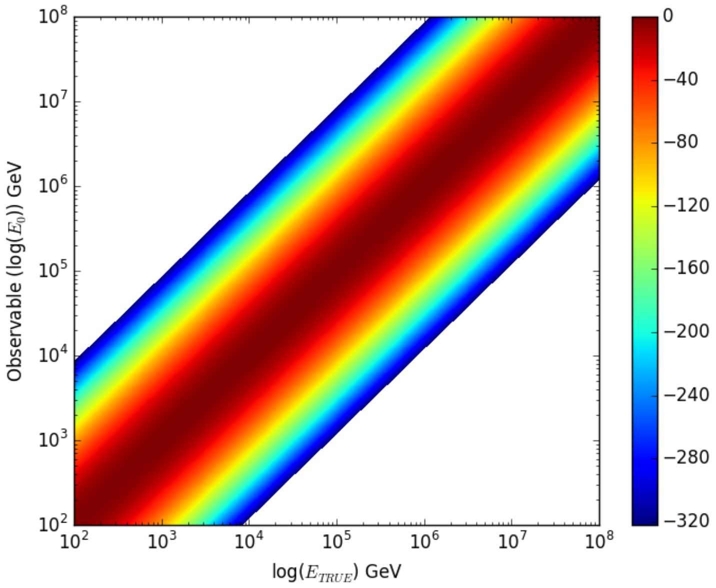


Figure 4.1: An example simulation of a Gaussian energy distribution for a muon energy reconstruction (E_0) with a true energy to reconstructed energy relationship denoted as $slope = 45^o$.

While Figure 4.1 allows for determination of the algorithm's accuracy, the spread of each true energy into reconstructed energy E_0 only provides information about one beam of true energies and is not a good indicator of overall resolution. For any given true muon energy there is a range of reconstructed energies that are possible; conversely any given reconstructed energy can correspond to multiple true energies.

There are two measures discussed in this chapter as methods to characterise resolution; these are the Average Likelihood Function and the Peak Likelihood. Both resolution measures can be used to compare energy reconstruction methods.

4.1.1 Average Likelihood Function Resolution

In the following section a method for displaying the resolution of the energy reconstruction distribution is described. This technique was also used to present the resolution of the truncated mean energy loss rate in [45].

The reconstruction distribution can be interpreted as if there was a beam of muons at a particular true energy, and then the slice vertically would be the distribution of reconstructed energies that those muons could be reconstructed as. However each reconstructed energy has a likelihood, taken as the horizontal slice of the plot, which gives the range of true energies that the given reconstructed energy could correspond to, i.e. the confidence interval of true energies given the observation. These confidence intervals need to be accounted for in the resolution by combining all the confidence intervals with a weight applied to each, corresponding to the probability of getting that reconstructed energy given the true energy of the muon beam. Overall, this results in an average likelihood function, for the target true energy. The one-sigma width is then taken as the resolution measure for that target true energy.

The reconstruction distribution in Figure 4.1 can be interpreted for a target true energy, $E_{TRUE}(j)$, along the reconstructed energy axis as the probability distribution function, $P(E_0; E_{TRUE}(j))$ of reconstructing the observable E_0 given the true energy E_{TRUE}. This is a slice at fixed E_{TRUE} taken vertically up the plot. For this distribution of possible observed reconstructed energies the probability distribution function represents a weight function such as shown in Figure 4.2 for $E_{TRUE} = 10^5$ GeV that can be applied to respective observed reconstructed energy likelihoods (horizontal slices).

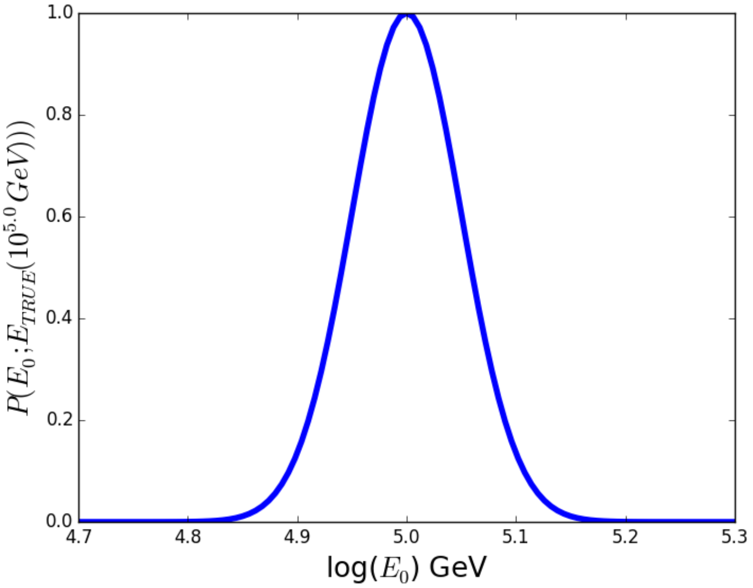
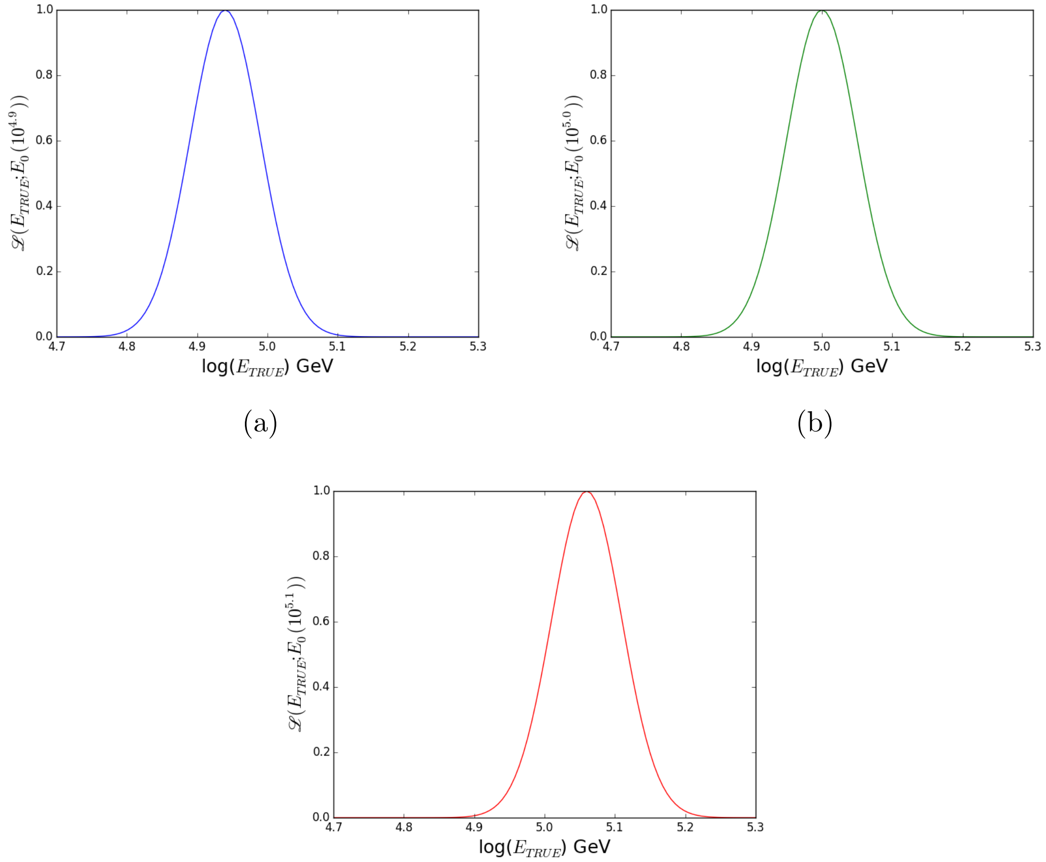


Figure 4.2: The probability for a target energy 10^5 GeV for the reconstructed distribution in Figure 4.1, used as a weight function in the resolution.

For a specific value of reconstructed energy the distribution along the true energy axis



(a)

(b)

(c)

Figure 4.3: The likelihood of a reconstructed observed energy corresponding to a true energy for (a) $E_0 = 10^{4.94}$ GeV (b) $E_0 = 10^5$ GeV (c) $E_0 = 10^{5.06}$ GeV

gives the likelihood of the true energy given fixed reconstructed energy E_0. As such it can be interpreted as the likelihood $\mathscr{L}(E_{TRUE}; E_0)$, of the parameter E_{TRUE} given a fixed observed energy E_0. The likelihoods for three observable energies are shown Figure 4.3, where each is seen to be centred around a different value of E_{TRUE}.

Each likelihood for a specific observable $E_0(i)$ in Figure 4.3 can be weighted by the probability from the weight function for the value of E_0, $P(E_0(i); E_{TRUE}(j))$ in Figure 4.2. Figure 4.4 shows the weighted likelihoods for the likelihoods in Figure 4.3, each weighted likelihood now has a different amplitude based on its weight value from Figure 4.2 for its observable energy E_0. This would be done on all energies and would result in a series of weighted likelihoods with varying amplitudes, as shown in Figure 4.5. The weighted likelihoods can be combined to make an average distribution as shown in Figure 4.6.

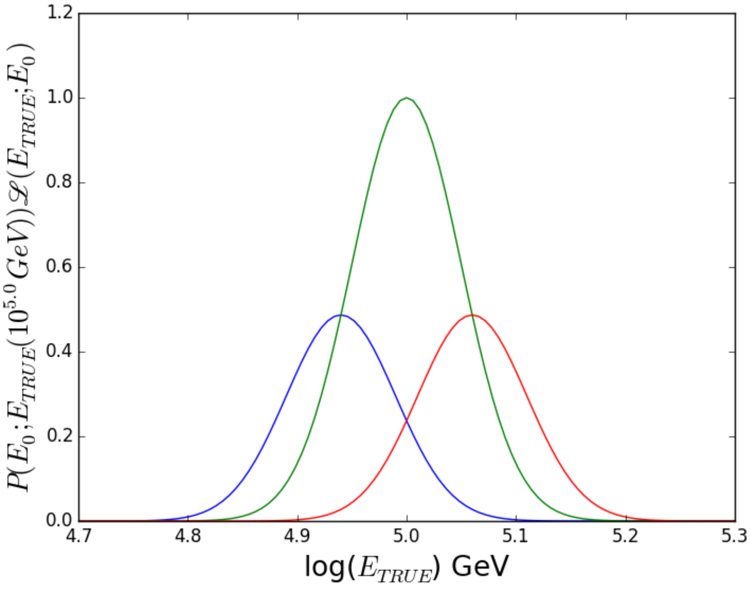


Figure 4.4: Three weighted likelihoods for a target energy 10^5 GeV for the reconstructed distribution in Figure 4.1 as calculated in Eq(4.1). Each colour represents an individual weight applied to each likelihood.

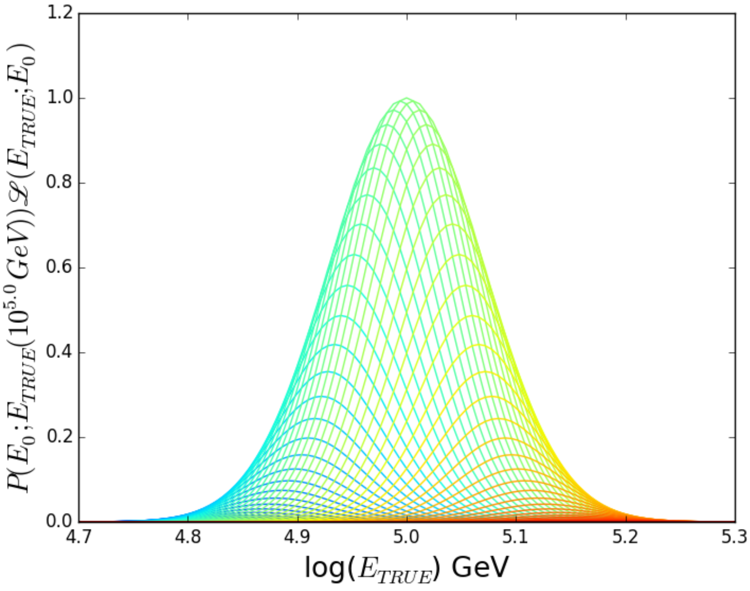


Figure 4.5: The series of weighted likelihoods for a target energy 10^5 GeV for the reconstructed distribution in Figure 4.1 as calculated in Eq(4.1). Each colour represents an individual weight applied to each likelihood.

The distribution in Figure 4.6 is the Average Likelihood Function (ALF) for the example target true energy, $E_{TARGET} = 10^5$ GeV. For any target true energy, we can write out the full expression for the weighted sum of the likelihoods as

$$P(E_{TARGET}; E_0) = \sum_{i=0}^{N} P(E_0(i); E_{TARGET}) \mathscr{L}(E_{TRUE}; E_0(i)); \tag{4.1}$$

where N is the number of energy bins along the observable energy axis and E_{TARGET} is the target true energy for which the distribution is found. In the example in this section the target energy was $E_{TARGET} = 10^5$ GeV, to find the resolution over all energies the procedure needs to be done at different target energies. The summed distribution calculated for a given true energy value would not be a distribution in true energy but rather is the distribution of the inferred muon energy.

The width of the distribution will give a measure (σ) of the reconstruction resolution. As a consistent measure of σ the width the value of the standard deviation for the distribution around each true energy will be taken and used to compare all reconstructions.

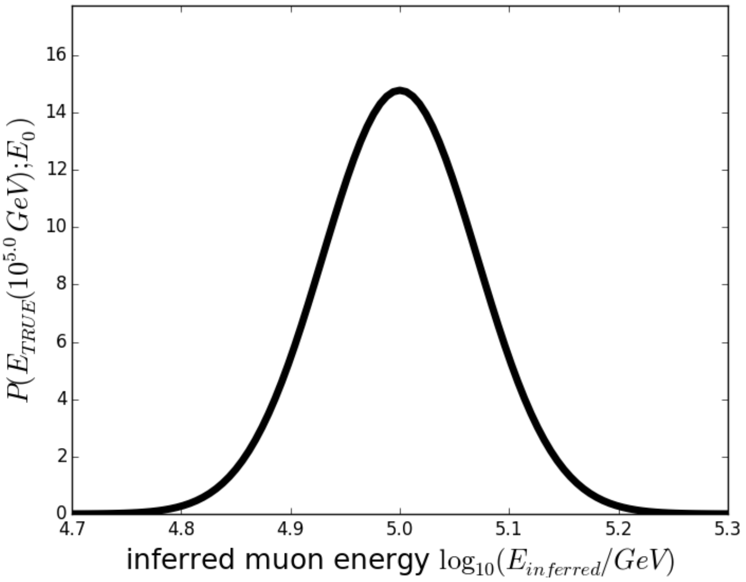


Figure 4.6: The ALF distribution for target energy 10^5 GeV along the inferred true muon energy, for the reconstructed distribution in Figure 4.1.

This method has the effect of mapping the reconstruction resolution distribution to be centred at the target true muon energy, thus removing systematic offsets, and it is able to show the resolution of the reconstruction distribution regardless of any offset in reconstructed and the true energies.

4.1.2 Peak Resolution Measure

An alternative is to treat the horizontal slices $\mathscr{L}(E_{TRUE}; E_0)$ not as distributions but to take the maximum or peak value of the distribution $\mathscr{L}(\hat{E}_{TRUE}; E_0)$ i.e. the value of true energy which is the maximum of the likelihood, $\hat{E}_{TRUE}$. The Peak measure of resolution follows the same procedure as Section 4.1.1, however it differs in its application of the observed energy likelihood function. Rather than use the whole function, only the energy at the peak in the likelihood $\mathscr{L}(\hat{E}_{TRUE}; E_0)$ is used, thus the resolution measure is represented by

$$P(E_{TARGET}; E_0) = \sum_{i=0}^{N} P(E_0(i); E_{TARGET}) \mathscr{L}(\hat{E}_{TRUE}(j); E_0(i)) \tag{4.2}$$

where N is the number of energy bins along the observable energy and E_{TARGET} is the target true energy for which the distribution is found. In this case $P(E_0(i); E_{TARGET})$ is the value of the weight function at that energy and $\mathscr{L}(\hat{E}_{TRUE}(j); E_0(i))$ is the peak value of the likelihood. This value corresponds to the most likely true energy for the particular observed energy. The series of weighted likelihoods that is produced is shown in Figure 4.7 with each point being a peak likelihood value that is weighted by the corresponding point on the weight function. The summed distribution given by Eq 4.2 is shown in Figure 4.8 for the target true energy value of $E_{TARGET} = 10^5$ GeV across the inferred muon energies.

4.1.3 Resolution Distribution Characterisation

In order to quote a standard measure of the resolutions for a distribution as in Figure 4.6 and 4.8 a value of the standard deviation for each distribution is taken. The value of standard deviation or σ is found via,

$$\sigma = \sqrt{\frac{\sum_{i=0}^{N} P(E_i)(E_i - \bar{x})^2}{\sum_{i=0}^{N} P(E_i)}} \tag{4.3}$$

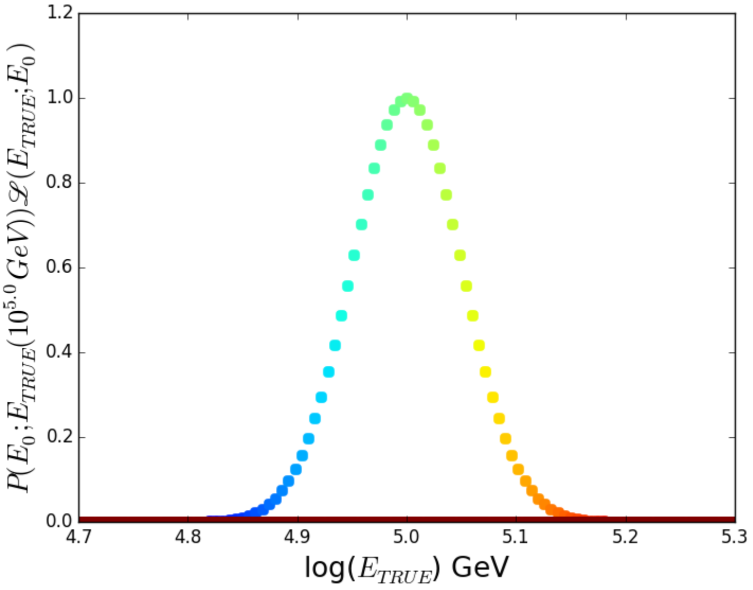


Figure 4.7: The weighted peak likelihoods for a target energy 10^5 GeV for the reconstructed distribution in Figure 4.1. Each colour represents an individual weight applied to the peak of the likelihood, thus there is only one value for each observable energy E_0.

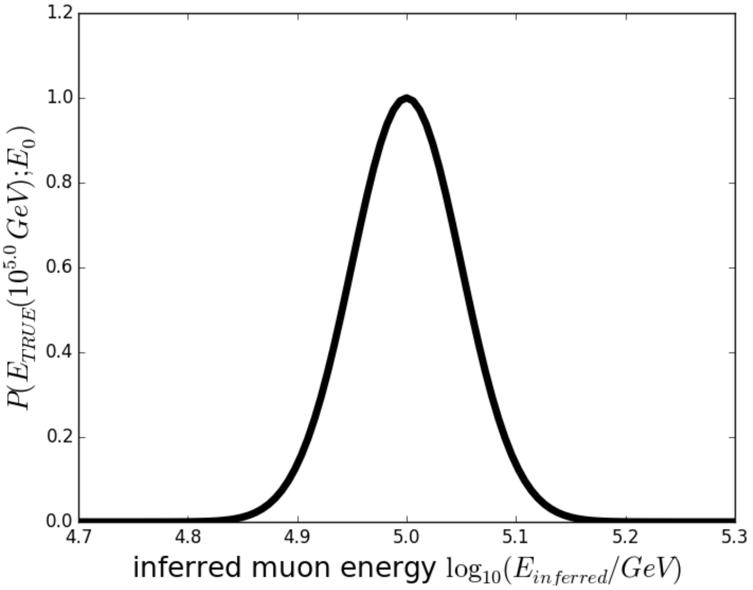


Figure 4.8: The Peak distribution for target energy 10^5 GeV along the inferred true muon energy, for the reconstructed distribution in Figure 4.1 calculated using Eq(4.2).

where $\sum_{i=0}^{N} P(E_i)$ is equal to the integral of probabilities in the distribution and provides normalisation to the distribution. The mean is given by

$$\bar{x} = \frac{\sum_{i=0}^{N} E_i P(E_i)}{\sum_{i=0}^{N} P(E_i)} \tag{4.4}$$

The comparison for ALF and Peak resolution distributions is shown in Figure 4.9 for $E_{TARGET} = 10^5$ GeV with the quoted σ's as $\sigma_{Peak} = 0.050$ and $\sigma_{ALF} = 0.071$.

While the distribution shows the level of reconstruction resolution for one energy it does not describe the resolution across all energies. This is fine in a idealised model where there is no variation in the distribution, but in the case of a realistic sample it is important to observe the changes in σ across a range of target energies, as reconstruction methods may perform better at certain energy ranges. For later chapters the variations in σ across energy ranges are calculated and used when defining the performance of the various reconstructions.

For optimal resolution it is desirable that the distribution be as narrow as possible, thus the smaller the value of σ the better the resolution. The ALF resolution distribution is by construction wider than the Peak resolution distribution due to the spread in the likelihood and the spread in the probability of the given true energy. In Figure 4.5 the separate weighted likelihoods are all symmetrical around the mean value, this creates a larger width at all values of inferred muon energy. This effect causes a slightly wider distribution for the ALF resolution than for the Peak resolution. However, both are correct measures of resolution and all further comparisons made in this work are relative comparisons using the same resolution measure. There is an additional effect in the Peak resolution distribution as the distribution retains the shape of the probability of reconstructed energy for true energy and thus experiences no smoothing effects of the distribution. This will be explored further in Section 4.2.2. Each measure describes different information that is valuable to understanding the reconstruction resolution and both are technically a correct measure of resolution. For the purposes of this work, where energy reconstruction methods are compared using their resolution, the same resolution measure is used for comparison to ensure consistent results. Due to its average functionality the ALF measure will be the standard shown in this work, in select cases, where applicable, both measures will be analysed.

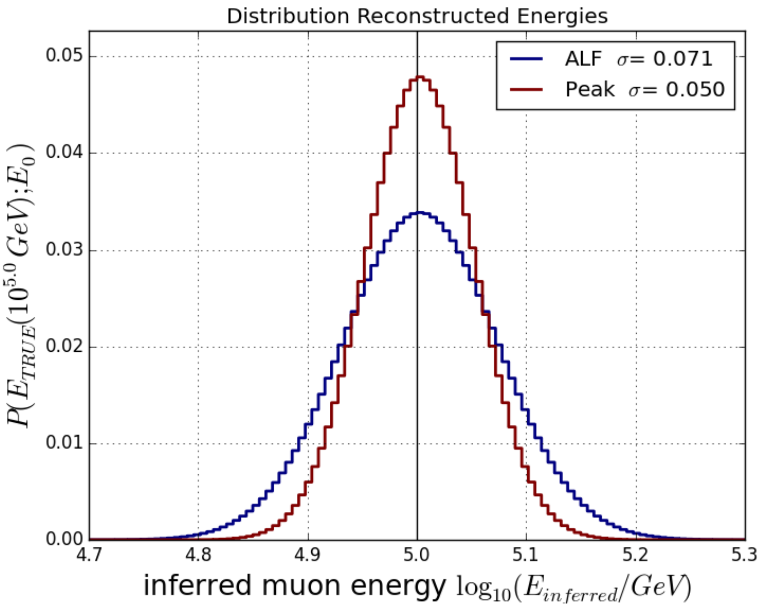


Figure 4.9: The resolution distributions for Figure 4.1 at $E_0 = 10^5\,\text{GeV}$ for the case of ALF resolution and Peak resolution

4.2 Distribution Effects on Reconstruction Resolution

4.2.1 Effect of Distribution Slope on Resolution

In this section first we will talk about the relationship between slope and width for the reconstruction distribution then we will provide an example of how the slope affects the resolution in this work.

The slope of the reconstruction distribution is an important factor in the reconstruction resolution. For example in the case of a near-horizontal distribution (where the reconstructed energy might change very slowly as true energy increases), while the width of the distribution along the reconstructed energy might be narrow for the true energy, it is not a good reconstruction as it is only able to reconstruct to that single observable energy. This is demonstrated in Figure 4.10 which shows a sketch of a reconstruction distribution typical of those seen in later chapters (e.g. Figure 6.45). The distribution has a change in slope at true energy E_B, where it goes from a horizontal distribution to a one-to-one distribution. It is important to note that the width along reconstructed energy for each true energy (vertical) is constant, thus this width on its own would not be a good measure of resolution. This is why resolution needs to account for the combination of the widths along reconstructed energy and the widths in reconstructed energy along true energy. The horizontal width along true energy is the range of true energies the reconstruction energy could correspond to, the smaller the range the better the reconstruction result. If the range is large there are more true energies the reconstructed energy could correspond to which makes it harder to distinguish between different true energies. In Figure 4.10 the widths along true energy for each reconstructed energy are different, for example at $E_0 = E_1$ (along the horizontal) the true energy width from E_A to E_B is larger than at $E_0 = E_2$ (along the one-to-one line) where the width in true energy is from E_C to E_D. The reconstruction resolution at E_2 will be better, as it has a smaller range of true energies, than the resolution at E_1, which has a larger range of true energies. There is another effect with the reconstruction distribution which will cause issues as all the true energies from E_A to E_B have been reconstructed to E_1, if the energy distribution was taken in reconstructed energy with counts, there would be a piling up of events at E_1 with no reconstructed energies below this energy. This would create a bias in the results that may affect an analysis, an example of this will be discussed later in this section.

In this section we will explore how each resolution measure accounts for the slope of the distribution. If the slope is not at 45^o there is an energy dependent bias that may affect the results. There are two limits to the slope of the reconstruction distribution at $slope = 0^o$ and $slope = 90^o$. At $slope = 90^o$ the distribution is vertical, meaning that regardless of the value of the observable it only corresponds to one true energy. This would of course only occur if the sample only contained one true muon energy bin. At $slope = 0^o$ the distribution is horizontal, such that the muon events are always being reconstructed to a single observable energy. In this case it would mean that the reconstruction is not working as there is no way of differentiating the reconstructed observables.

A resolution measure should be able to identify the problems with these limiting cases. The resolution distribution for the horizontal case for the ALF resolution results in a uniform resolution distribution across all energies. The value of σ for these cases will be large and so will clearly show the problems in the distribution. In the Peak case the

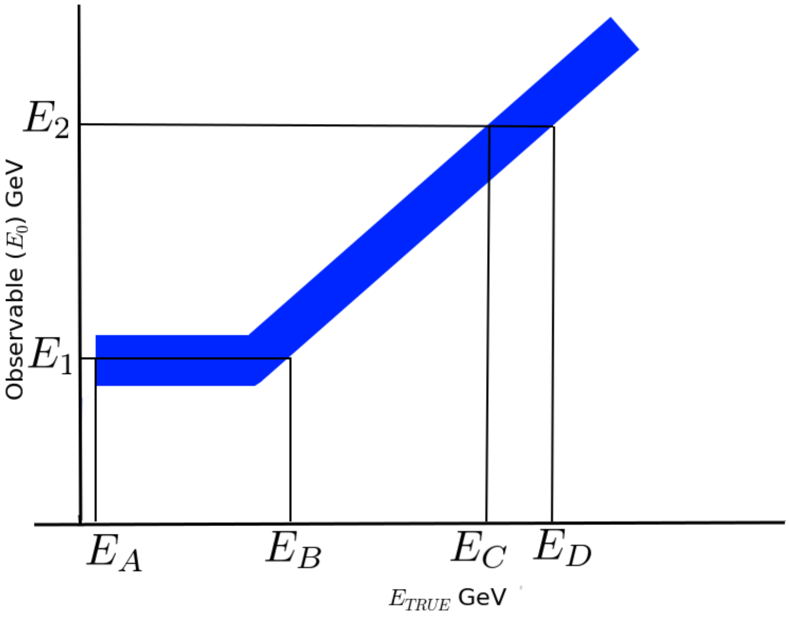


Figure 4.10: A sketch of a reconstruction distribution with true energy E_{TRUE} vs the reconstructed energy E_0, with a change in slope, this sketch is similar to what is seen in Figure 6.45. The reconstruction has the same width in reconstructed energy for each true energy (vertical). At E_1 the reconstruction distribution has a slope of 0^o and thus will only reconstruct to E_1, for all true energies E_A to E_B. However, at E_2 there is a one-to-one slope which will have a better resolution than that of E_1, as there are a smaller range of true energies the reconstructed energy could correspond to from E_C to E_D.

value of resolution would be $\sigma = 0.0$ which would indicate a non-functioning resolution.

For the vertical limit the resolution distribution only occurs at one value of E_{TRUE}. There is no clearly defined error in this distribution as it does give a resolution for this true energy but any study of σ across all energies would identify this problem.

Between the limiting cases of slope, there is a range of possible distributions. The reconstruction can be defined as working as long as the distribution is not one of the two limiting cases, ie in cases where $0^o < slope < 90^o$.

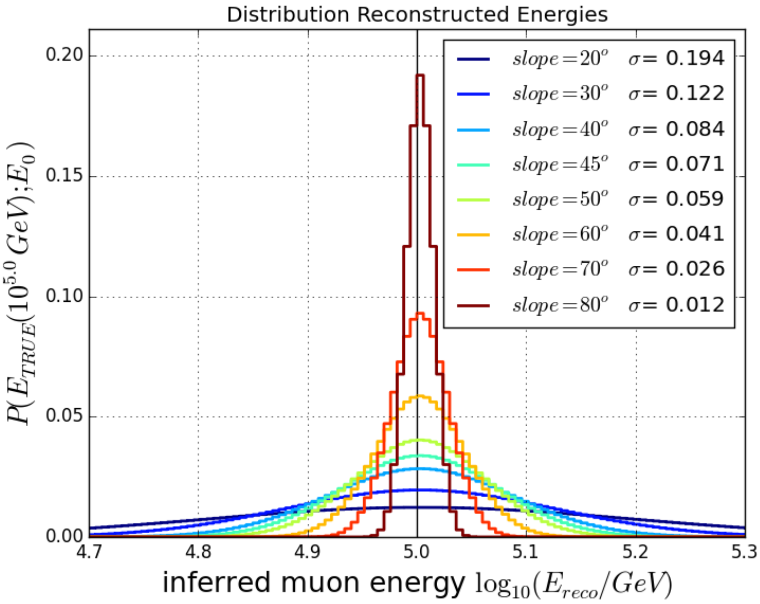


Figure 4.11: Reconstruction resolution using the ALF measure for examples with increasing slope in the relationship between E_0 and E_{TRUE}. The spread in each reconstruction energy is fixed at $\sigma = 0.05$.

The resolution results of various distribution slopes shown in Figure 4.11 and 4.12 demonstrate the inverse relationship between resolution and slope; as the slope increases, but keeping the spreads in reconstructed energy constant for each true energy, the σ_{ALF} decreases, signifying a better resolution for the reconstruction distribution. As the probability weight function is the same for both resolution measures the changing slope has the same effect on both the Peak and ALF resolutions.

Due to the relationship of resolution with respect to slope the increasing slope on the distribution will continue to produce a better resolution as it approaches $slope = 90^o$. However it is important to note that a vertical reconstruction distribution of $slope = 90^o$

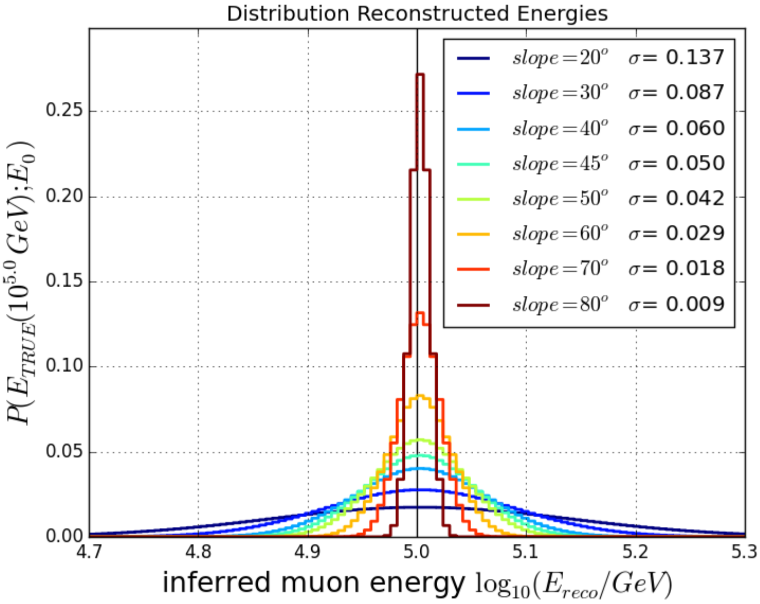


Figure 4.12: Reconstruction resolution using the Peak measure for examples with increasing slope in the relationship between E_0 and E_{TRUE}. The spread in each reconstruction energy is fixed at $\sigma = 0.05$.

is not a working reconstruction: the reconstruction is unable to distinguish between true energies. Therefore as long as $slope < 90^o$ the reconstruction is functional, provided the observable energy has binning to cover the desired energy range.

The effect of slope on resolution influences the results in later chapters, the sketch in Figure 4.10 is based on the reconstruction distribution in Figure 6.45 where the slope is flatter at lower energies resulting in all true energies below 10^3 GeV being reconstructed in the range $10^3 - 10^4$ GeV. This has an effect in the diffuse analysis in Chapter 8 where in Figure 8.4 the reconstruction distribution causes an offset in the convolved distribution of energies, with events piling up at energies just above 10^3 GeV which is the reconstructed energy corresponding to the low slope part of Figure 6.45.

4.2.2 Resolution of Abnormal Reconstruction Distributions

Any measure of reconstruction resolution needs to be able to adapt and encompass abnormal reconstruction distributions as may potentially occur. Such abnormal reconstruction distributions may be physical or are an error in the reconstruction that needs to be identified. The resolution measure should be equipped to interpret such distributions for analysis and diagnostics.

An example abnormal distribution with a double peaked Gaussian structure in the probability of E_0 given E_{TRUE} is shown in Figure 4.13 for all true energies. This abnormal distribution produces the reconstruction resolution distributions in Figure 4.16. The separate measures produce different shapes in the resolution distributions due to the differences in their construction.

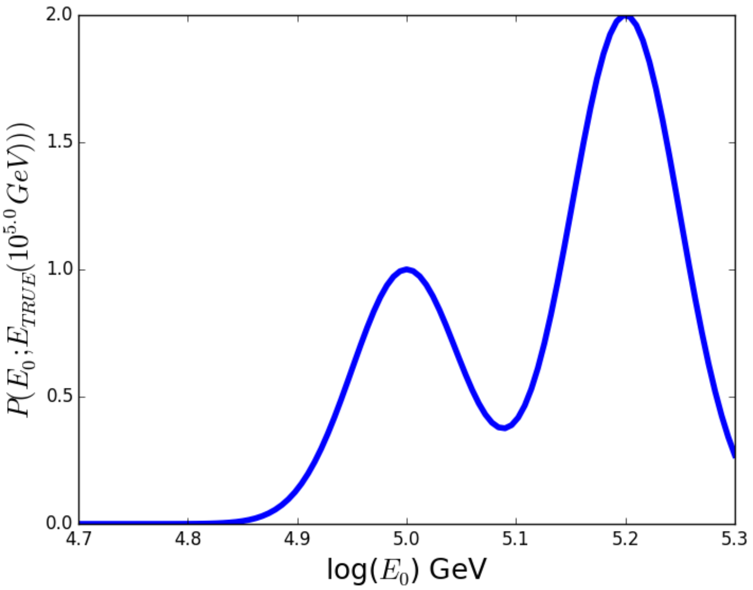


Figure 4.13: The weight function corresponding to 10^5 GeV for the abnormal double maximum distribution. The weight function shows the double peak characteristic of the distribution which, while not expected in a typical reconstruction, may be produced in error.

The individual weighted likelihoods in Figure 4.14 and 4.15 show how the Peak and ALF measures differ because of the treatment of the likelihoods. In the case of the Peak measure in Figure 4.15 the distribution maintains the same shape as in the reconstruction distribution along the true energy. This is due to the nature of the method which causes the probability to be multiplied by the peak values which, for a correctly normalised distribution, are a series of constant values. As such the distribution is a reproduction of the shape of the probability with its amplitude adjusted by a fixed value of the peak.

Using the ALF measure a series of likelihoods are taken into account. Each likelihood

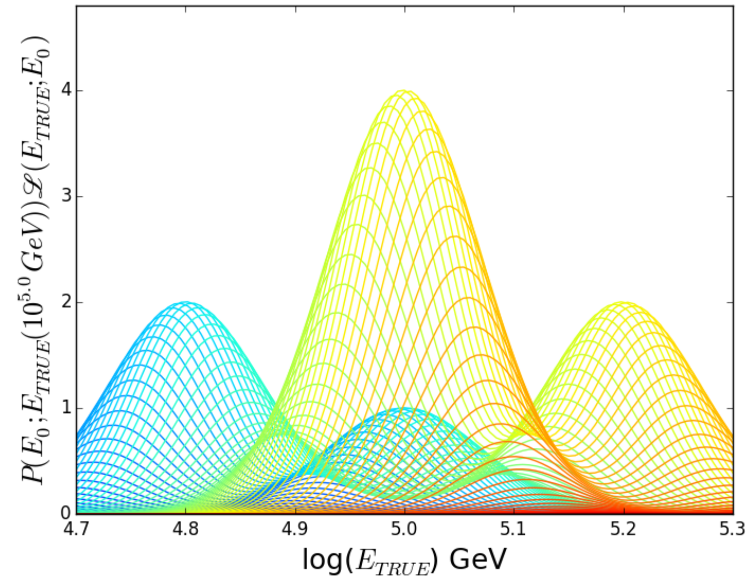


Figure 4.14: The weighted likelihoods for the target energy 10^5 GeV for the abnormal example in Figure 4.13 for the case of ALF resolution.

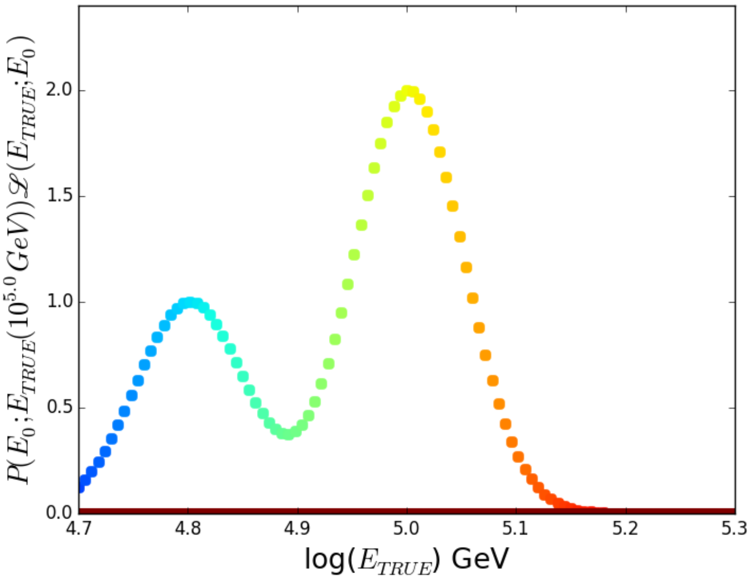


Figure 4.15: The weighted likelihoods for the target energy 10^5 GeV for the abnormal example in Figure 4.13 for the case of Peak resolution.

has a weight that causes an increase as the target energy is approached, from both higher and lower energies. This means along the edges of the distributions there are likelihoods with low weight which still contribute a full likelihood shape which cause wider wings. In the case of the abnormal distribution this is more apparent as the use of the full likelihood causes a bump on both sides of the target energy. This is a double counting effect that can be clearly seen in Figure 4.14 which causes a wider distribution. However taking the average in the ALF measure has the effect of smoothing the resolution distribution which makes for a clearer value of σ.

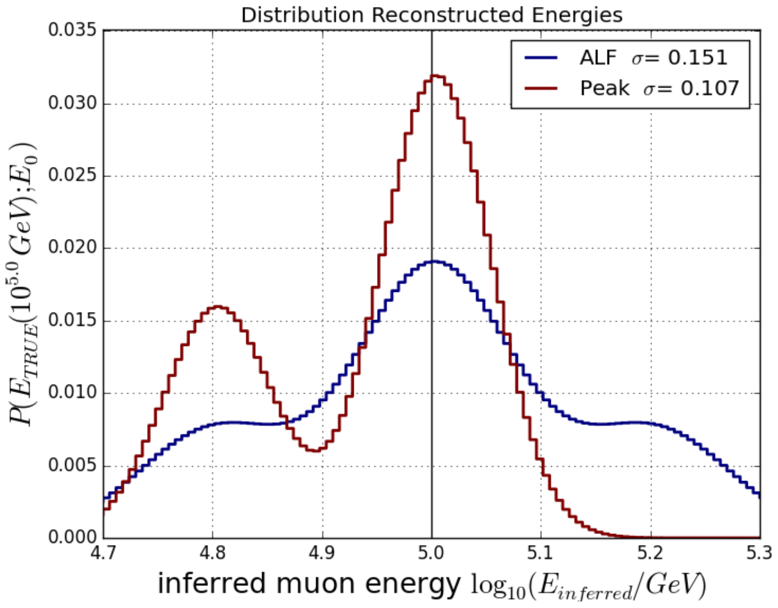


Figure 4.16: Distribution of reconstruction resolution for the abnormal example with the probability weight function given in Figure 4.13 for the case of ALF resolution and Peak resolution.

4.3 Technical Details of Reconstruction Resolution

The measures of resolution are dependent on having an accurate distribution of the likelihoods and probabilities for the target energies. In the ideal example given in Section 4.1 and 4.2 this is not a problem, but when these measures are used in later chapters on muon simulation various statistical effects need to be accounted for.

In later chapters, measures of resolution are taken from the reconstruction of muon events that are in a binned reconstruction distribution. The value of the sample size in the distribution i.e. the number of statistics in the sample, obviously can have an effect on the accuracy of the measure of the resolution. With too few statistics, fluctuations occur in the resolution measure for different target true energies. The energy reconstructed distributions in this work have a minimum of 10^4 muon events to achieve an adequate resolution measure, and where possible a greater simulation sample size was used.

There is also a limitation in the resolution measure for distributions at the edges of the energy range. At the edges of the energy range the probability of reconstructed energy given true energy would not be complete and thus would appear to have a smaller width. Additionally if the entire likelihood is not contained within the energy range the average will not be accurate and the peak measure will not be able to find the maximum likelihood value. Due to this the incomplete probability weight and likelihood gives the impression of the reconstruction energy being more accurate than is the case. This would appear in a resolution energy range plot as a downturn in the resolution measure at the edges of the energy range. These edge effects are present in both resolution methods as both utilise the probability of a reconstructed energy at a given true energy. In this work the solution to minimise the edge effects is to cut the resolution results in energy such that only results obtained with complete probability and likelihood distributions are used. In all future resolution plots the energy range will only show results that comply with this requirement.

Finally, the Peak measure relies on the maximum likelihood occurring in monotonically increasing true energy bins. If this is not the case, and instead the likelihood's peak value occurs in the same true energy bin as the next observed energy's likelihood, the summation to inferred muon energy will result in the addition of energy bins to create artificial peaks in the distribution with adjacent empty bins, such as shown in Figure 4.17. This would not allow for correct calculation of the standard deviation. This error could occur in any reconstruction as any slight variation of the peak from the mean will result in an empty energy bin.

This limitation with the selection of the peak is something not experienced by the ALF measure, as taking the average over the full likelihood naturally avoids this. In order to produce Peak resolution distributions a fit can be performed for the peak values between

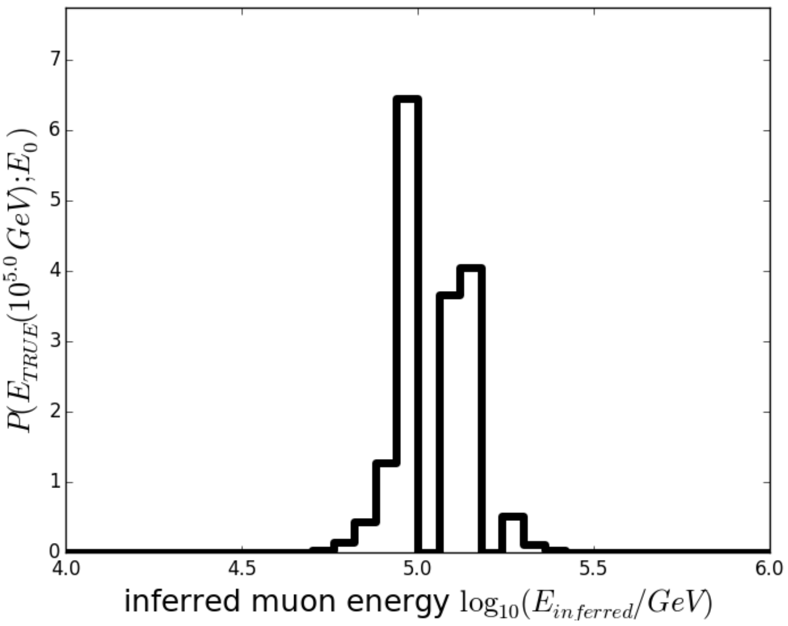


Figure 4.17: The Peak resolution distribution for target energy 10^5 GeV for a Gaussian reconstructed distribution.

true energy and observable energy. For the results in Figure 4.18 the peak bins have been selected and a first order polynomial fit to map the peak in reconstructed energy vs true energy has been applied. This approximation can be used to place a peak value in every successive energy bin, this then results in the complete distribution seen in Figure 4.18.

The requirements for finding an accurate measure of reconstruction resolution discussed here have been implemented for all following chapters. This allows the resolution measure to be used as consistent measure of reconstruction performance.

4.4 Summary

The energy resolution of a reconstruction method can be characterised by its reconstruction resolution, which can be measured by the ALF or Peak method. Both methods use the likelihood of a reconstructed loss over true energies combined with the probability distribution of reconstructed energies given a target true energy. The ALF method utilises the whole likelihood function to create an average for the resolution distribution that will have a continuous distribution. The Peak method only uses the peak of the likelihood

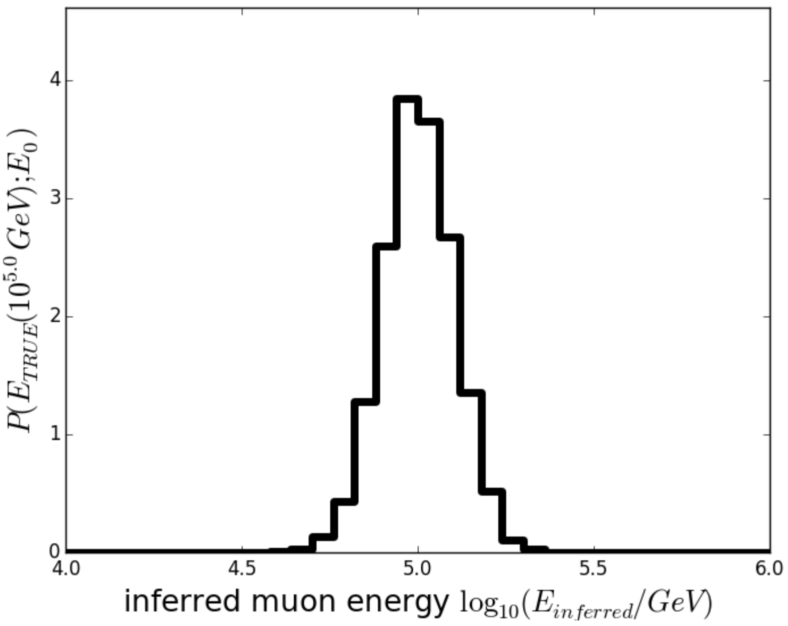


Figure 4.18: The Peak resolution distribution for target energy 10^5 GeV for a Gaussian reconstructed distribution with an approximation included for the peak likelihood values.

which creates a narrower resolution distribution but has limitations in the implementation that sometimes require a line to be fitted to the peak values.

While the Peak method gives the smaller measure of resolution, both methods are valid for assessing the resolution of an energy reconstruction as long as a relative comparison is done using the same resolution measure (ALF or Peak).

Chapter 5

Edepillim: Muon Energy Reconstruction Method

This work is based around the development of a new muon energy reconstruction method that, unlike previous methods, uses the whole energy loss pattern of the muon. This new method capitalises on the stochastic nature of muon interactions to produce an energy reconstruction that we will show in the idealised case has better energy resolution than any previous method when using the true energy loss pattern (e.g. truncated energy (TE), discussed in Section 3.4).

The initial concept for this method was based on using the energy loss pattern produced in the event reconstruction algorithm Millipede and working backwards from the pattern to get to the original energy. Thus the method was named the semi-palindrome; Edepillim. The following chapter describes the Edepillim method implementation and performance on simulated muon events. We will also test the effect of basic detector limitations on the Edepillim method by varying the energy loss segment length and muon track path length.

5.1 The Edepillim Method

The Edepillim method is a topological muon reconstruction that uses the probabilities of each observed muon energy loss within the energy loss pattern to achieve a high resolution energy reconstruction result [1].

A travelling muon will undergo stochastic energy loss processes as described in Section 3.1, which will result in an uneven pattern of energy loss along the muon track. The amplitude of each energy loss will be related to the muon's energy immediately prior to the loss. In general high energy muons will have larger energy losses than those of low energy muons.

This idea can be expressed in terms of a probability likelihood function for a muon with an initial energy E_0 and an energy loss pattern $d\vec{E} = [dE_0, dE_1, dE_2, ..., dE_N]$, where dE_0 is the first energy loss, dE_1 is the next energy loss in the pattern and so on until energy loss dE_N. The likelihood of the total observed loss pattern is based on successive probability distribution functions for each individual loss:

$$P(d\vec{E}; E_0) = P(dE_0; E_0)P(dE_1; E_1)...P(dE_N; E_N) \tag{5.1}$$

where there are N losses and each $P(dE_i; E_i)$ represents an individual probability of seeing an energy loss dE_i , over a segment length X, for a particular muon energy E_i.

The probability in Eq 5.1 can be simplified, as the muon energy after the first energy loss, E_1, is simply the difference between the initial muon energy, E_0, and the previous energy loss, dE_0,

$$E_1 = E_0 - dE_0 \tag{5.2}$$

The same principle applies to the energy after the next loss,

$$E_2 = E_0 - dE_0 - dE_1 \tag{5.3}$$

The progression of this along the muon track means the energies of the muon at any particular step can be expressed in terms of the initial muon energy and the sum of the previous energy losses,

$$E_N = E_0 - \sum_{i=0}^{N-1} dE_i \tag{5.4}$$

---

[1] This type of muon energy reconstruction method was developed by Petrukhin and Kokoulin in [47]. This method was brought to our attention by Petrukhin at a poster presentation of the Edepillim results during its development.

The relationship in Eq 5.4 can be visualised in the sketch in Figure 5.1 of an energy loss pattern with the energy at each point along the pattern related to the energy losses that came previously.

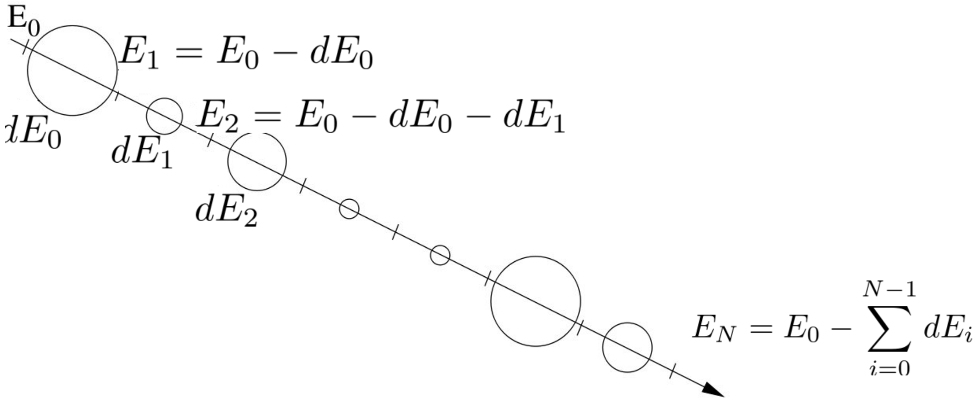


Figure 5.1: A sketch of the muon energy loss pattern. The amplitude of each loss is given by the radius of the circle. The energy loss at any point can be given by Eq 5.4.

Because of this, successive muon energies can, for each probability, be put in terms of the original muon energy E_0. Thus a probability likelihood function can be written as the following

$$P(d\vec{E}; E_0) = P(dE_0; E_0) \prod_{j=1}^{N} P(dE_j; E_0 - \sum_{i=0}^{j-1} dE_j) \tag{5.5}$$

Knowing the energy losses of the muon pattern and having a detailed probability distribution function of muon energy losses, a maximum likelihood approach to Eq 5.5 can be used to compute the unknown variable; the muon's initial energy E_0.

In the rest of the chapter we explore the implementation of this method on simulated true energy losses and examine the energy resolution the method obtains. The simulation was done using PROPOSAL to propagate muons along a path length, thus results in an idealised case of using Edepillim reconstruction on true energy losses. The detector conditions taken into consideration were a total muon path length of 1 km and variations on possible segment lengths of the energy losses, to demonstrate the effect of a detector energy loss resolution on the final reconstruction result.

5.1.1 Probability Distribution Function

In order to reconstruct an accurate muon energy the Edepillim method requires an accurate Probability Distribution Function (PDF). In this section we describe the production of the segment energy loss PDF $P(dE;E)$ and the fractional energy loss PDF $P(V;E)$.

Segment Energy Loss PDF

The PDF needs to provide the probability of an energy loss occurring given a muon energy ($P(dE;E)$), however in practice not every individual energy loss can be known due to detector resolution. Instead, a total of all the energy losses would be known in a given segment length X, thus each PDF would need to be made for a required segment length $P_X(dE,E)$. To build the PDF the simulation software PROPOSAL will generate a muon at energy E and allow it to travel along a segment length X. Then the energy loss is taken as $dE = E_{initial} - E_{final}$, where $E_{initial}$ is the energy the muon was generated at and E_{final} being the generated muon's energy after a segment length X. Taking the losses in terms of total energy rather than the option of using the generated list of energy losses allows for the continuous energy losses to also be accounted for, as not all of these losses are in the list of energy losses.

Muons are generated at 800 energies between 10^2 GeV $-$ 10^{10} GeV along a log scale. In order to obtain high statistics for each muon energy with segment length X, 100,000 iterations of the muon propagation are done with that initial energy.

The PDF results shown in Figure 5.2 are for a segment length of 1 m and the PDF is normalised such that

$$\int P(dE;E)d(dE) = 1 \tag{5.6}$$

for each muon energy E. The PDF in Figure 5.2 displays features that may have effects on the performance of the reconstruction. The diagonal one-to-one line is the physical cut off for energy losses not to be greater than the initial energy of the muon. This is an important physical constraint to enforce in the reconstruction because if you have a muon with a given energy loss you must only consider muon energies greater than the given loss.

The PDF's vertical line represents the minimum value of energy loss for the given

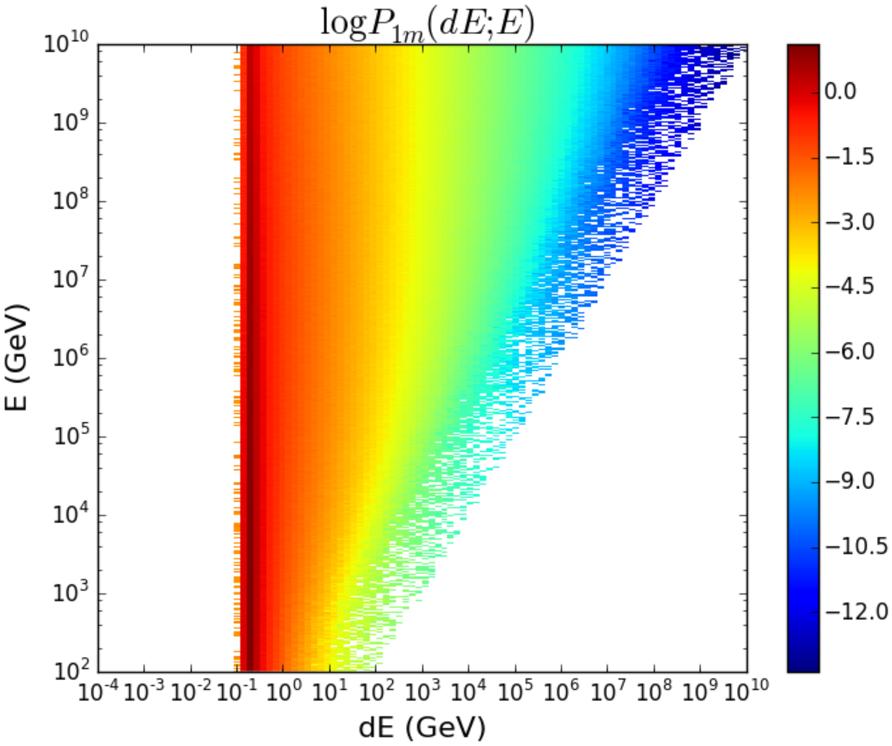


Figure 5.2: Probability distribution function built from simulated muon energy losses within a segment length $X = 1\,\mathrm{m}$, with each probability taking into account the bin width in dE, to convert the total counts per bin into a PDF, displayed in the log of probability. There is a diagonal line which represents the limit of the energy loss being greater than the muon energy. The vertical line is the minimum probable energy loss $dE = 0.2\,\mathrm{GeV}$.

segment length X, and, as continuous losses are included, there is a minimum of approximately 200 MeV/m from ionisation. The use of continuous losses in the PDF impacts the implementation of Edepillim as all losses in the true or reconstructed energy loss pattern will need to have a continuous component to them. In the idealised case these continuous losses are included, but in Chapter 6 the Millipede reconstructed energy loss patterns do not always contain continuous losses and thus care must be taken in defining the PDF.

For each segment length of a muon that will be used a unique PDF needs to be produced. In Figures 5.3 to 5.6, for a range of segment length values, the PDFs show similar features to Figure 5.2 but as the segment length increases, such as in Figure 5.4 where $X = 50\,\mathrm{m}$, the minimum probable energy loss increases to 1 GeV. This is due to the continuous loss being proportional to segment length. There is also an effect at the higher muon energies; as the segment length increases the minimum probable energy loss does not remain at the ionisation energy loss but moves to a higher energy loss. In Figure 5.5 this can be observed as above $E = 10^4\,\mathrm{GeV}$ there is another diagonal line representing the lower limit of the probable energy loss. When the segment length is $X = 500\,\mathrm{m}$ in Figure 5.6, the PDF is only a small diagonal band. This energy loss cutoff at higher energies is due to the nature of the cross-section of muon interactions for large segment lengths where multiple interactions are more likely to occur.

Fractional Energy Loss PDF

There is an alternative form of the $P(dE; E)$ PDF to that shown in Figures 5.2 to 5.6, which is to bin in fractional energy losses $V = dE/E$ instead of the direct energy losses dE. This was found to be in some cases a simpler implementation method and in other cases resulted in problems with the likelihood space. Both versions of the PDF are mathematically equivalent if an appropriate weighting factor is used in the likelihood, (see Section 5.1.2). Both forms will be used within this research with the distinction made between $P(dE; E)$ and $P(V; E)$.

The $P(V; E)$ PDF is produced in the same manner as $P(dE|E)$ but normalised according to

$$\int P(V; E)dV = 1 \tag{5.7}$$

for each muon energy E.

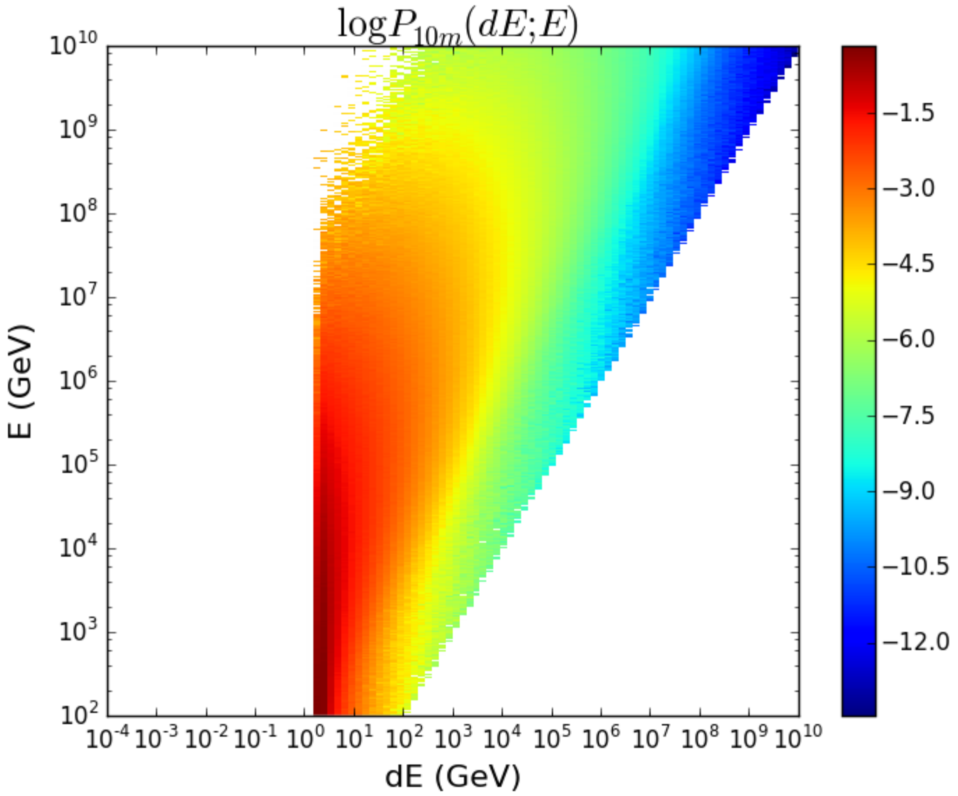


Figure 5.3: Probability distribution function built from simulated muon energy losses within a segment length $X = 10\,\mathrm{m}$, with each probability taking into account the bin width, to convert the total counts per bin into a PDF, displayed in the log of probability. There is a diagonal line which represents the limit of the energy loss being greater than the muon energy. The vertical line is the minimum probable energy loss $dE = 2\,\mathrm{GeV}$. From approximately $E = 10^{8}\,\mathrm{GeV}$ the minimum energy is less probable as there are more stochastic losses causing higher energy losses to be more probable.

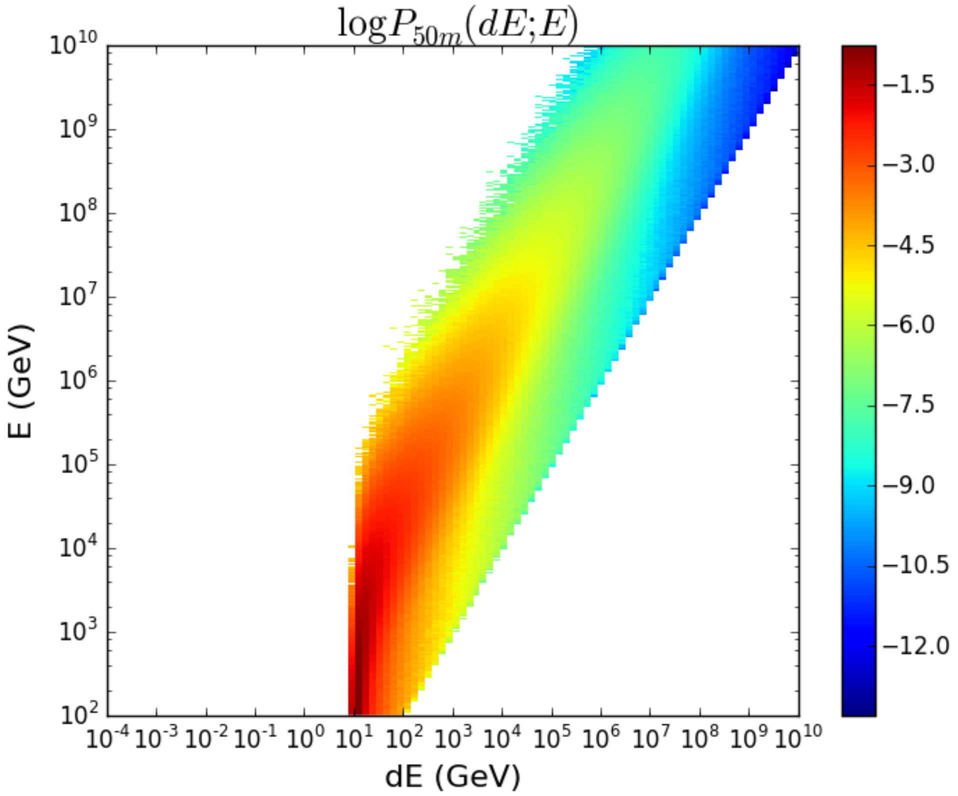


Figure 5.4: Probability distribution function built from simulated muon energy losses within a segment length $X = 50\,\mathrm{m}$, with each probability taking into account the bin width, to convert the total counts per bin into a PDF, displayed in the log of probability. There is a diagonal line which represents the limit of the energy loss being greater than the muon energy. The vertical line is the minimum probable energy loss $dE = 10\,\mathrm{GeV}$. From approximately $E = 10^5\,\mathrm{GeV}$ the minimum energy is less probable as there are more stochastic losses causing higher energy losses to be more probable, this occurs at lower energies than in Figure 5.3 as the larger segment length means more energy losses are present.

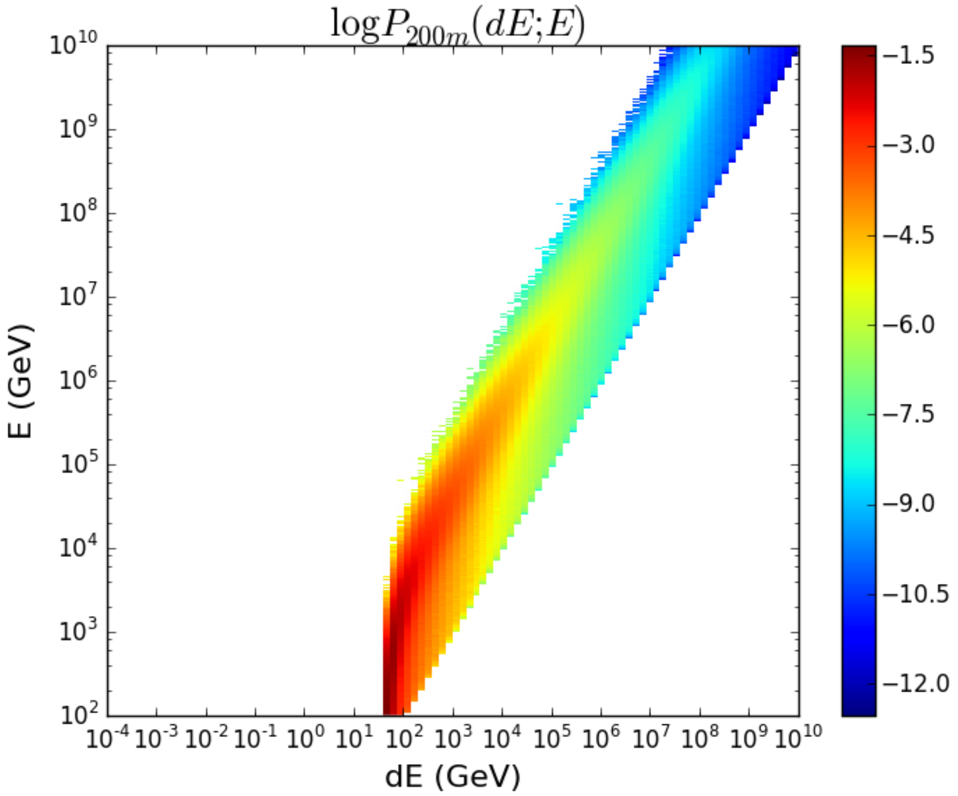


Figure 5.5: Probability distribution function built from simulated muon energy losses within a segment length $X = 200\,\mathrm{m}$, with each probability taking into account the bin width, to convert the total counts per bin into a PDF, displayed in the log of probability. There is a diagonal line which represents the limit of the energy loss being greater than the muon energy. The vertical line is the minimum probable energy loss $dE = 40\,\mathrm{GeV}$. From approximately $E = 10^4\,\mathrm{GeV}$ the minimum energy is less probable as there are more stochastic losses causing higher energy losses to be more probable. These two limits result in the PDF being only a small band along the range of energy losses.

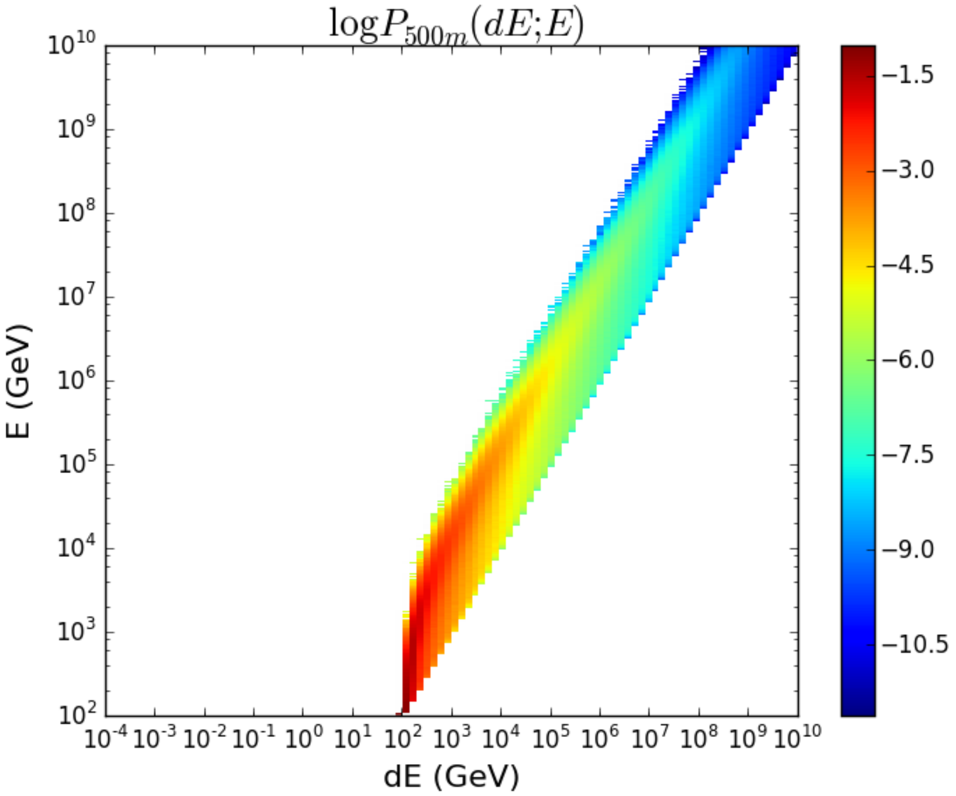


Figure 5.6: Probability distribution function built from simulated muon energy losses within a segment length $X = 500\,\mathrm{m}$, with each probability taking into account the bin width, to convert the total counts per bin into a PDF, displayed in the log of probability. There is a diagonal line which represents the limit of the energy loss being greater than the muon energy. The vertical line is the minimum probable energy loss $dE = 100\,\mathrm{GeV}$. From approximately $E = 10^4\,\mathrm{GeV}$ the minimum energy is less probable as there are more stochastic losses causing higher energy losses to be more probable. As in Figure 5.5 the two limits result in the PDF being only a small band along the range of energy losses.

Figure 5.7 shows the PDF for a 1 m segment length, which displays the same physical constraints as Figure 5.2. The minimum energy loss at 200 MeV/m appears as a diagonal line representing $V_{min} = 0.2/E$ for each muon energy E(GeV). There is also a physical limitation at $V = 1$ as the energy loss cannot be greater than the muon energy.

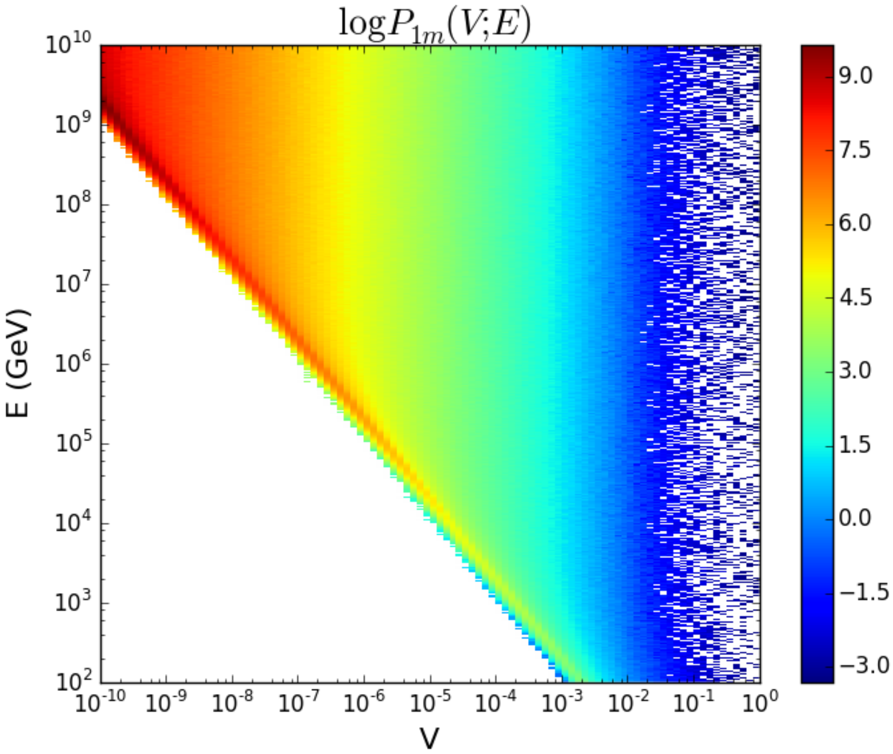


Figure 5.7: Probability distribution function built from simulated losses of segment length $X = 1$ m binned along fractional energy loss $V = E/dE$, displayed in the log of probability. The minimum energy loss produced by PROPOSAL causes the diagonal cut off at increasing V per decreasing muon energy E.

5.1.2 Maximum Likelihood Scan

In this section we describe the process of the Edepillim methods maximum likelihood scan for both the segment energy loss and the fractional energy loss implementation.

Segment Energy Loss Implementation

To implement a maximum likelihood search the Edepillim method performs a grid scan over a range of initial muon test energies in the likelihood space. This range is based on how far the PDF extends in E as for anything outside this range we would have no physical information about.

Certain physical limitations are imposed on the reconstruction, the major one being that there is no point considering a test energy that is below the sum of the energy losses as a muon cannot produce more total energy losses ($\sum dE$) than its initial energy. This motivates the different starting energy for the likelihood scan in different muon events, as the sum of the energy losses represents a lower bound on the possible reconstructed muon energy.

Another limitation is that in the true energy loss case if an event has an energy loss of zero at the end of the energy loss pattern the physical effect is that the event has lost all of its energy. As a result the reconstructed energy is set to the sum of all the energy losses in the pattern as the muon must have had this total energy as it started depositing energy along its path length in the detector. This effect is dominant with muons below 10^3 GeV, which tend to range out over hundreds of metres. Such a limitation can be imposed in the case of a true simulation as all energy losses are true losses but should not be imposed on the case of reconstructed losses due to the uncertainty in the accuracy of the reconstructed losses as explored in Chapter 6 (where we show that Millipede reconstructs an abundance of zero energy losses). Additionally, as a range of initial test energies is tried there will be cases where the PDF for a particular energy loss will return a zero probability. This is due to the binned nature of the PDFs, the simulated event may have an energy loss that corresponds to an empty bin in the PDF. For example in Figure 5.6 any test energy above 10^4 GeV will return a zero probability for a minimum energy loss of 100 GeV, thus those test energies should not be allowed in the likelihood space as such a energy loss for the test energy should not occur. In order to avoid this, any case that returns a probability of zero will be disallowed in the likelihood space by rejecting that test energy as a possible minimum. This results in a smaller range of test energies that are possible.

The maximum likelihood is performed using the log likelihood of Eq 5.5 giving

$$\ln L = \sum_{i=0}^{N} \ln P(dE_i; E_i) \tag{5.8}$$

where dE_i is the energy loss in a bin i and E_i is the muon energy at the start of the bin i and N is the total number of energy loss segments along the path length.

As the likelihood scans over test initial muon energies, for each point in the scan there is a starting test energy E_{test}, which will be tried as E_0. The probability for the first energy loss is given by the PDF as $P(dE_0; E_0)$. For the next bin, $i = 1$, the decreased muon energy is now taken as the current energy: the difference of the test initial energy and the energy loss i.e. $E_{current} = E_0 - dE_0$. This process continues for all bins of the energy loss pattern giving a total log likelihood value.

The process is then repeated for successive test energies. Provided no physical limitations are reached each test muon energy will be covered during the grid scan. The grid scan over test energies produces a likelihood space such as shown in Figure 5.8 with a negative log likelihood used, thus there will be a minimum. This minimum becomes our observable E_0, the reconstructed energy, which for this event is $E_0 = 5.6 \times 10^4$ GeV, which is close to the true energy of the muon $E_{TRUE} = 4.2 \times 10^4$ GeV.

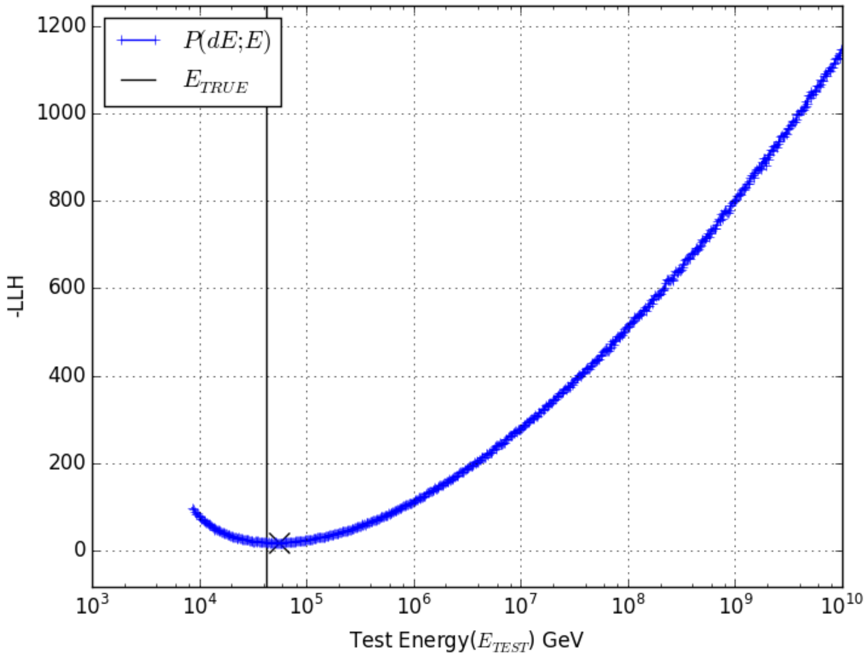


Figure 5.8: An Edepillim likelihood space for a particular muon event of energy $E_{TRUE} = 4.2 \times 10^4$ GeV; the cross shows the reconstructed energy $E_0 = 5.6 \times 10^4$ GeV that was found, the vertical line shows the true energy for the muon. The cut off to the left of the likelihood space is the physical limitation due to the sum of the energy losses, $\sum dE = 8.6 \times 10^3$ GeV, which sets a lower bound to the possible initial muon energies.

The range of muon energies the Edepillim method is able to reconstruct at is limited by the range of energies that the PDF was generated over. For this work the minimum test energy used was $E = 10^2\,\mathrm{GeV}$ based on limits in the production of the PDF. It is also limited by the sum of the energy losses as discussed earlier in this section.

Fractional Energy Loss Implementation

As discussed in the Section 5.1.1 there is an alternative form of the PDF using the fractional energy loss ($P(V;E)$). To implement this PDF in a maximum likelihood function we first need to understand the relationship between the two PDFs. Both forms of the PDF are made using the same method, it is merely in the binning along the x-axis (dE or V) that the differences manifest. As such both PDFs at any specific energy $E = E_s$ would normalise to the same value as given by Eq 5.6 and 5.7 [48]

$$\int P(dE;E_s)d(dE) = \int P(V;E_s)dV = 1 \tag{5.9}$$

The PDF used for the Edepillim method is not a functional form. It is a binned PDF, and as such the integral can be approximated as

$$\sum_{j}^{N_{bin}} P(dE_j;E_s)\Delta(dE_j) = \sum_{j}^{N_{bin}} P(V_j;E_s)\Delta(V_j) = 1 \tag{5.10}$$

where N_{bin} is the bin number along dE and V respectively and $\Delta(dE_j)$ and $\Delta(V_j)$ are the widths of the bins. Eq 5.10 can be simplified by substituting the fractional energy loss

$$\sum_{j}^{N_{bin}} P(dE_j;E_s)\Delta(dE_j) = \sum_{j}^{N_{bin}} P(V_j;E_s)\Delta(dE_j/E_s) \tag{5.11}$$

which as E_s is a constant leads to the relationship between the two PDFs

$$\sum_{j}^{N_{bin}} P(dE_j;E_s) = \frac{1}{E_s}\sum_{j}^{N_{bin}} P(V_j;E_s) \tag{5.12}$$

This relates the PDFs over all bins in the x-axis but the individual bins in each PDF can also be related. If you have the binning for dE in the range (dE_1, dE_2) where all values dE_s will be entered into one bin then in V the range would be $(dE_1/E_s, dE_2/E_s)$ which as E_s is a constant would mean all values that were entered into a bin in dE would be in the corresponding bin in V. Thus the probability for an energy loss for the two PDFs would be equivalent for each individual bin via

$$P(dE; E_s) = \frac{1}{E_s} P(V; E_s) \tag{5.13}$$

Thus the two PDFs merely differ by a factor of E_s and the relationship can be generalised for all energies. Therefore the PDFs can be considered equivalent in the maximum likelihood scan so long as this factor is taken into account.

To use the fractional energy loss implementation the maximum likelihood would substitute the PDFs as in Eq 5.13 to yield:

$$\ln L = \sum_i^N \ln P(dE_i; E_i) = \sum_i^N \ln \frac{1}{E_i} P(V_i; E_i) \tag{5.14}$$

If the likelihood equation does not account for this relationship it will not be a functioning maximum likelihood. If instead, one leaves out the $1/E$ term the following was taken as the likelihood function

$$\ln L = \sum_i^N P(V_i; E_i) \tag{5.15}$$

then the relationship of Eq 5.13 can still be applied term-by-term giving

$$\ln L = \sum_i^N \ln E_i P(dE_i; E_i) \tag{5.16}$$

$$\ln L = \sum_i^N \ln E_i + \sum_i^N \ln P(dE_i; E_i) \tag{5.17}$$

The final likelihood function has an additional factor of $\sum \ln E_i$ that is not in Eq 5.8. This factor will not be a constant across the likelihood space. As the test energy increases $\sum \ln E_i$ will increase. In the likelihood space this will have the effect of a constantly decreasing likelihood value that overpowers the true minimum.

Using the weight factor $1/E$ the alternative fractional implementation of the method becomes mathematically equivalent to the original energy loss implementation.

5.2 Edepillim Energy Reconstruction Results in an Idealised Case

In order to test the Edepillim reconstruction method a simple simulation of muons were produced using PROPOSAL to propagate a muon of a initial energy over a path length of 1000 m with the energy losses reported for every segment length $X = 1$ m. For the purpose of testing the performance of Edepillim under detector like conditions a simulation was produced of 800,000 muon events in the energy range 10^2 GeV to 10^{10} GeV, as described in Appendix A.

The energy loss patterns produced in this idealised simulation, such as shown in Figure 5.9 with initial muon energy $E_{TRUE} = 10^4$ GeV, have a continuous loss at the minimum ionisation loss rate as well as larger stochastic losses that are produced due to the processes of Bremsstrahlung, pair production and photonuclear reactions. The energy loss patterns are consistent between events with the amplitude of the stochastic losses scaling with the muon's energy, see Figure 5.10 where the initial muon energy was $E_{TRUE} = 10^9$ GeV which has energy losses orders of magnitude higher than Figure 5.9. The one variation in energy loss patterns is for low energy muons such as in Figure 5.11 where the muon has run out of energy just before 500 m of propagation. In such cases the reconstruction will determine that the muon has deposited all of its energy along the path and thus the initial energy at the entry must have been the sum of all the energy losses along the subsequent path length.

5.2.1 Muon Energy Loss Segment Length

In a detector of kilometre scale, accurate energy loss pattern reconstruction can be limited depending on the medium and the detector. In a detector such as IceCube the ice can cause scattering of the photons making the reconstruction of the location of the energy losses uncertain. For a detector such as ANTARES the water medium has a larger scattering length and thus the energy loss pattern can be estimated with more certainty in the segment length [49]. Because of this possible uncertainty in the location and size of reconstructed energy losses it is unlikely that the 1 m energy loss resolution as is used in

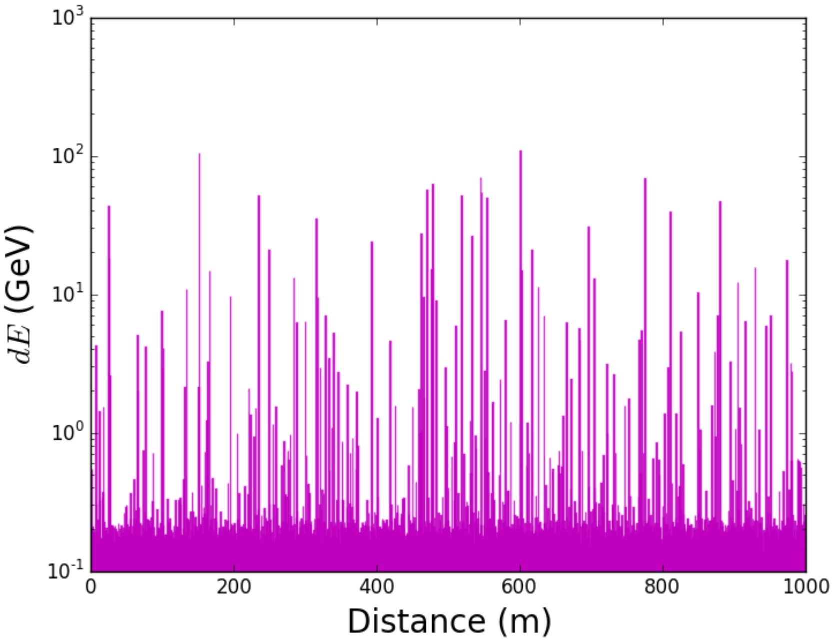


Figure 5.9: An example of a energy loss pattern for a muon event simulated at $E_{TRUE} = 10^4$ GeV with energy losses in a segment length $X = 1\,\mathrm{m}$. There is a base line of continuous losses with larger stochastic losses also produced along the path.

this simulation would be possible. To show the performance of Edepillim under different detector conditions the individual energy loss patterns of the simulated events were combined to make energy loss segment lengths of 10 m, 20 m, 50 m, 100 m, 200 m and 500 m. The purpose of larger segment lengths is to show how the performance of Edepillim changes as the expected resolution of the reconstructed energy loss pattern changes. The energy loss pattern's change with an $X = 10\,\mathrm{m}$ segment length is shown in Figure 5.12; the energy losses are still stochastic with additional continuous losses. In Figure 5.13 when the segment length is $X = 100\,\mathrm{m}$ the energy losses are more uniform with no clear distinction for stochastic energy losses.

The performance of the Edepillim energy reconstruction for segment length $X = 1\,\mathrm{m}$ is shown in Figure 5.14 with the reconstruction distribution having a consistent peak that correlates to the true muon energy, E_{TRUE}, and the reconstructed muon energy, E_0. With the increase in segment length, the reconstruction distributions become wider but maintain a consistent peak, such as in Figure 5.15, this continues for larger segment lengths to distributions such as Figure 5.16. In Figure 5.17 when using $X = 500\,\mathrm{m}$ which

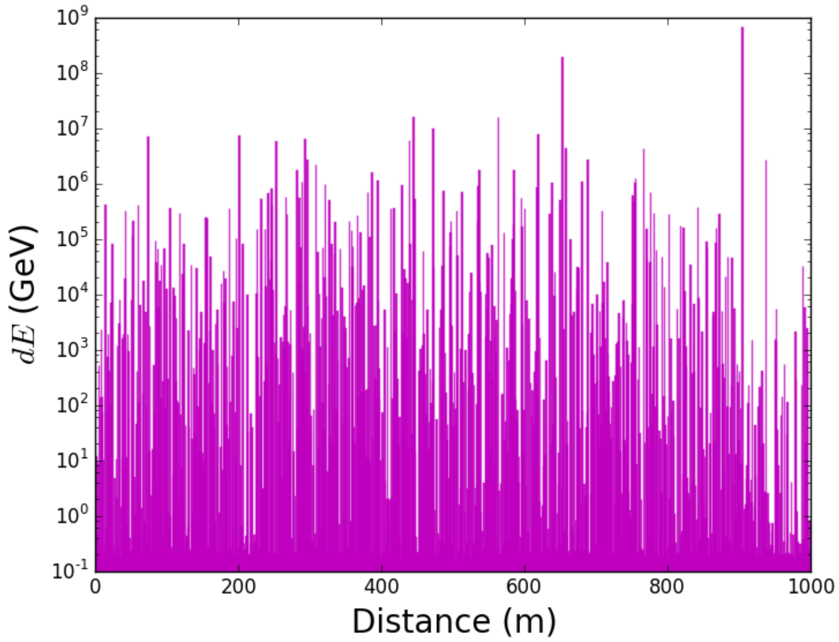


Figure 5.10: An example of a energy loss pattern for a muon event simulated at $E_{TRUE} = 10^9\,\text{GeV}$ with energy losses in a segment length $X = 1\,\text{m}$. This is an example of a high muon energy. The energy loss pattern is more dominated by stochastic energy losses than that seen in Figure 5.9.

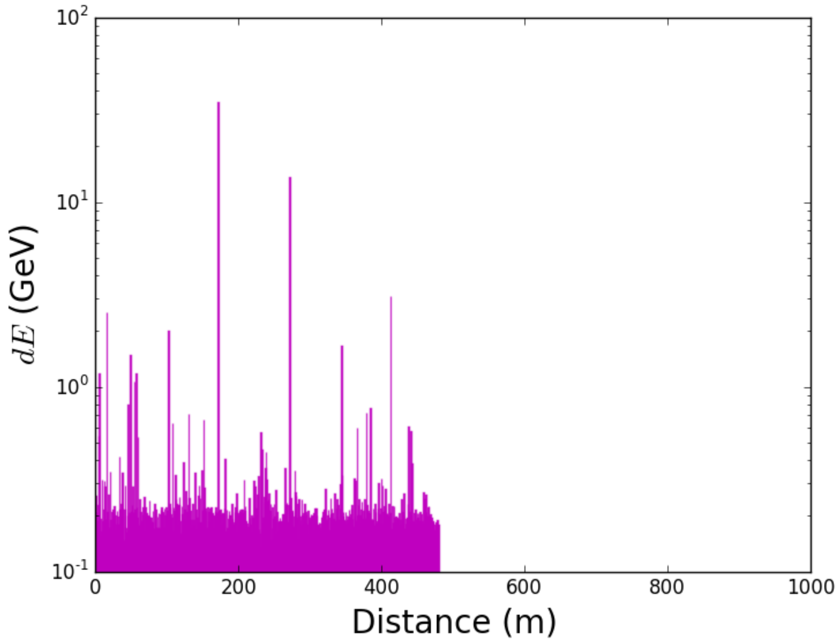


Figure 5.11: An example of a energy loss pattern for a muon event simulated at $E_{TRUE} = 10^2$ GeV with energy losses in a segment length $X = 1$ m. The low energy of the muon results in its shortened path length as the muon loses all of its energy about 500 m into the path.

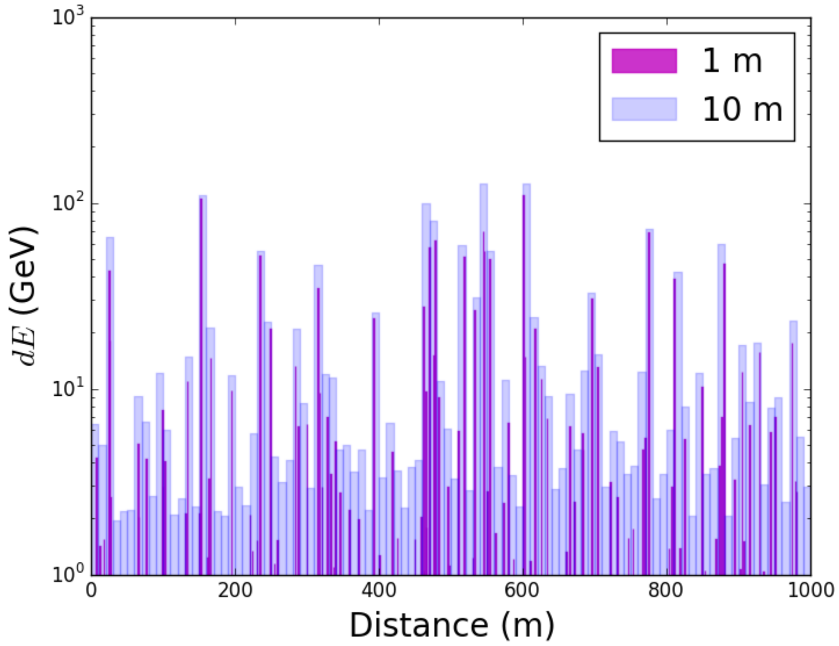


Figure 5.12: An example of a energy loss pattern for a muon event simulated at $E_{TRUE} = 10^4$ GeV with energy losses in a segment length $X = 1$ m combined into a segment length of $X = 10$ m. The combined segment length results in individual bins having higher energy loss but with the stochastic energy losses still being distinguishable.

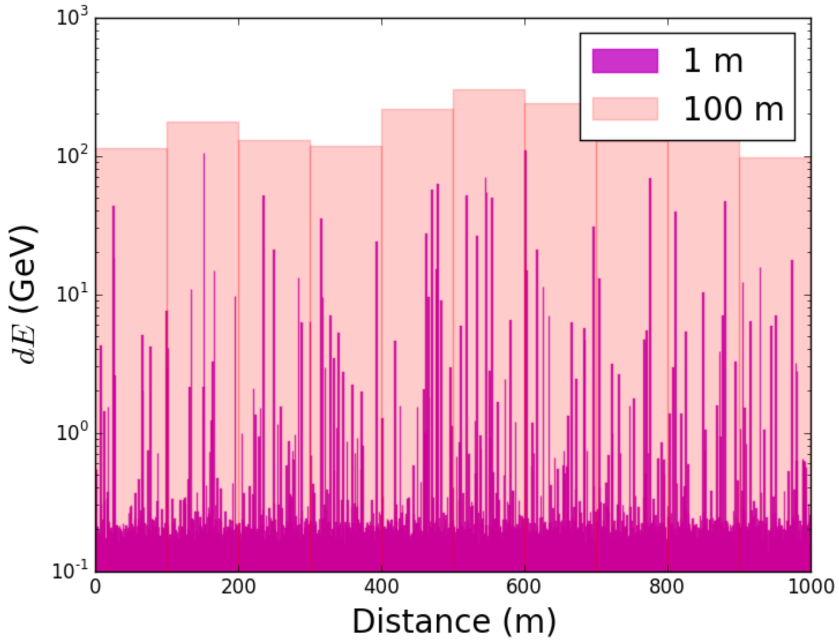


Figure 5.13: An example of a energy loss pattern for a muon event simulated at $E_{TRUE} = 10^4\,\mathrm{GeV}$ with energy losses in a segment length $X = 1\,\mathrm{m}$ combined into a segment length of $X = 100\,\mathrm{m}$. The combined segment length results in a more uniform energy loss pattern with difficulty distinguishing the stochastic energy losses.

would be only two losses, the reconstruction distribution has a clear smearing effect across true energies.

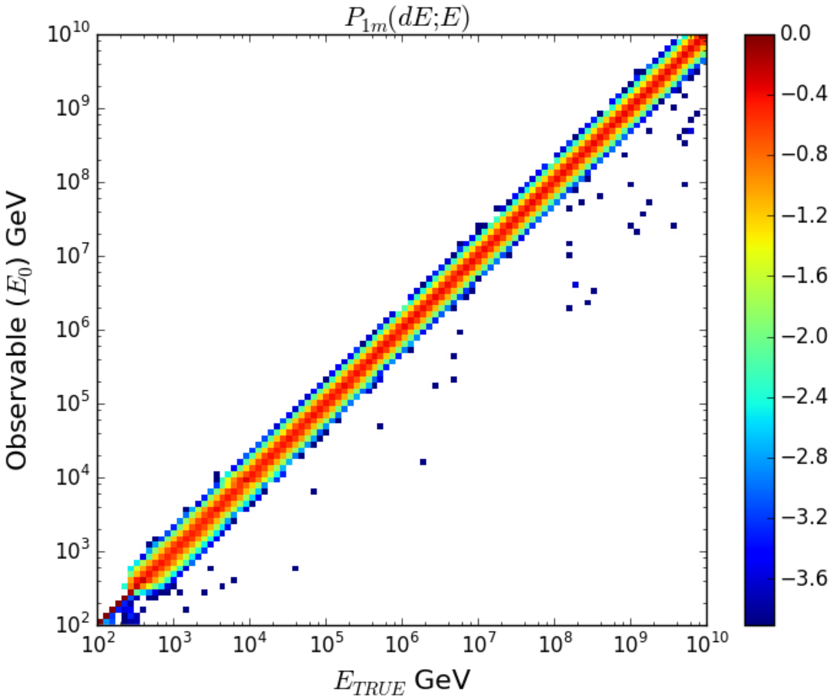


Figure 5.14: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 1\,\text{m}$ and total path length $L = 1000\,\text{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}.

To understand the performance of the Edepillim reconstruction method a measure of resolution is calculated using the resolution distribution as described in Chapter 4. Figure 5.18 and Figure 5.19 show the resolution obtained from the respective reconstructions, using the ALF and Peak measures across a range of energies.

The Edepillim method has optimal resolution with energy losses binned in $X = 1\,\text{m}$ intervals. As the segment length increases the individual energy losses get grouped together and so the detail in the energy loss pattern is lost. This reduces the Edepillim reconstruction performance as there will be fewer probabilities in the likelihood with increasing segment length. In Figure 5.18 the larger binned energy loss lengths result in a poorer resolution. However it is interesting to note that in the $X = 500$ m case in Figure 5.17 there are only two energy losses, and while this is not enough information for the

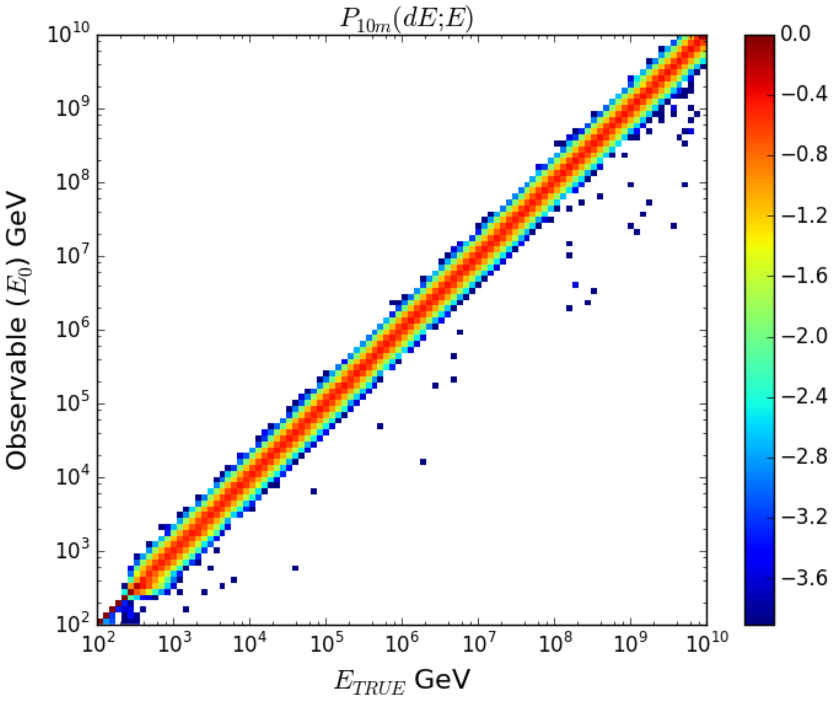


Figure 5.15: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 10\,\mathrm{m}$ and total path length $L = 1000\,\mathrm{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}.

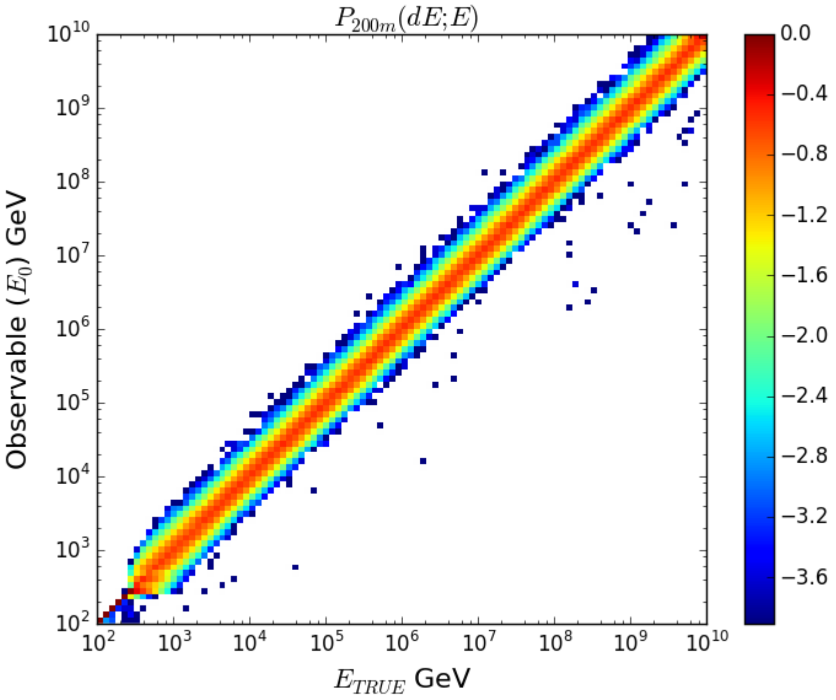


Figure 5.16: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length 200 m and total path length $L = 1000$ m. Displayed in the logarithm of muon events, normalised per E_{TRUE}.

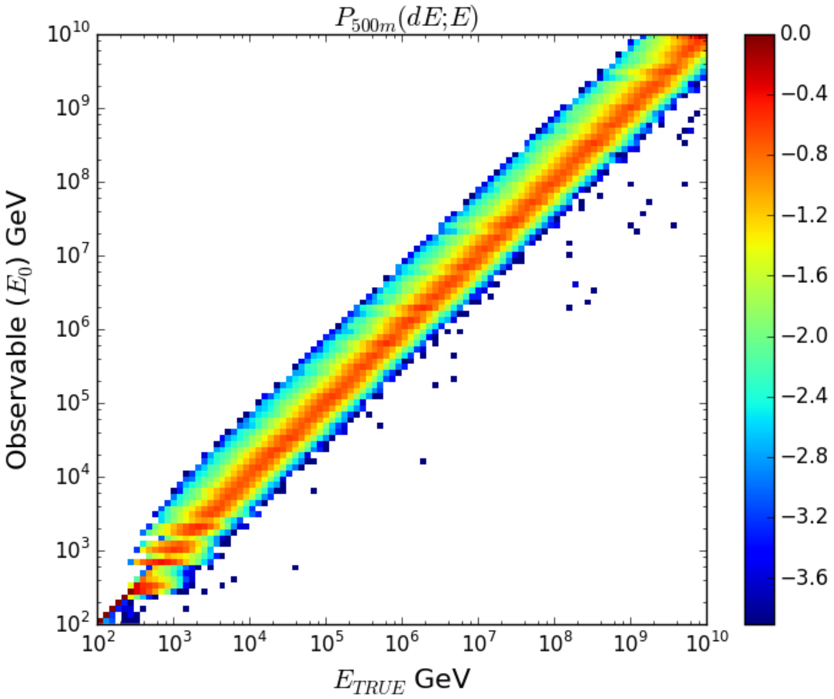


Figure 5.17: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 500\,\mathrm{m}$ and total path length $L = 1000\,\mathrm{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}.

Edepillim method to perform with high resolution, it is still a working reconstruction, despite only two input dE values. Once the number of losses doubles in the $X = 200$ m rebinned case (Figure 5.16) the resolution in Figure 5.18 improves significantly with a large gap in resolution results. The decreasing segment lengths follow the same trend but with smaller differences in the resolution results.

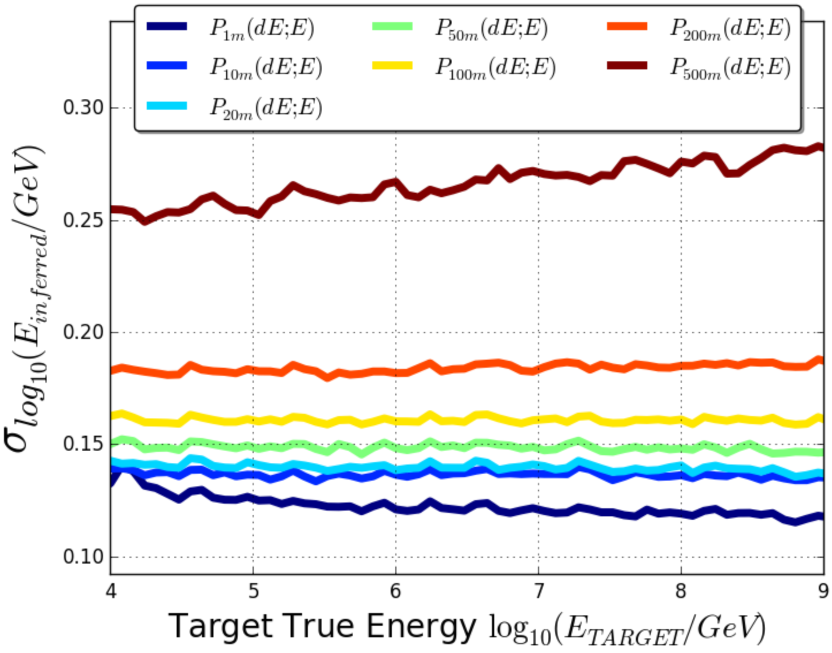


Figure 5.18: The Edepillim energy reconstruction ALF resolution measure for respective energy loss segment length over a range of target true muon energies. The reconstruction resolution increases as the energy loss segment length increases, indicating the best resolution if for energy loss patterns with the smallest segment length.

The results in Figure 5.18 and Figure 5.19 show how the optimal resolution for energy reconstruction can be obtained. The level of resolution achieved in a realistic case with real reconstruction of the losses will depend on how accurately the energy loss pattern can be known. If we know the losses at a 1 m scale, then the energy reconstruction can be obtained with a resolution of $\sigma = 0.12$ in the logarithm of the target true energy.

5.2.2 Muon Energy Loss Pattern Path Length

In kilometre scale detectors such as IceCube not all muon events will travel in a path that allows for a 1000 m energy loss pattern to be reconstructed. Through-going muons will

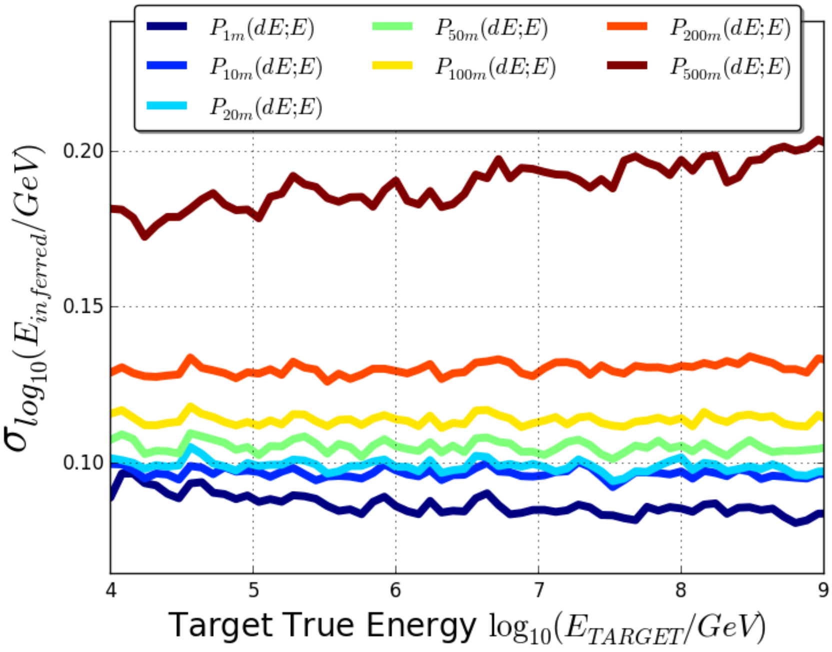


Figure 5.19: The Edepillim energy reconstruction Peak resolution measure for respective energy loss segment length over a range of target true muon energies. The Peak resolution follows the same trend as Figure 5.18 with the reconstruction resolution increases as the energy loss segment length increases, indicating the best resolution if for energy loss patterns with the smallest segment length.

more likely pass though at angles that do not pass through the centre of the detector and thus would have a path length of less than 1000 m. Also in very rare cases muons may be produced inside the detector, giving a starting track.

To simulate the potential cases of shorter muon tracks the muon energy loss patterns were limited to path lengths of 750 m, 500 m and 250 m. These path lengths do not mean that the muon has run out of energy, merely that the shorter length is all that is observed of the muon. The results for the case of $X = 1\,\mathrm{m}$ segment length with muon track lengths less than 1000 m in Figure 5.23 displays the large loss in resolution power of the reconstruction with shortened track lengths. The impact on the reconstruction is seen in the resolution in Figure 5.20 to Figure 5.22. There is no change in the slope of the reconstruction distribution but the width of the resolution becomes wider, with it becoming noticeable for path lengths 250 m and 500 m in Figures 5.22 and 5.21. It is also important to note the change in resolution at lower energies. In Figure 5.20 the events below 10^4 GeV are still well reconstructed but as the path lengths decreases to Figure 5.22 there is a more smeared reconstruction distribution and there is no longer a consistent peak line, which will result in effects in the reconstruction resolution.

These results show that the total muon path length has a large impact in the performance of the reconstruction. The effect is even greater than that shown by energy loss segment length, for as soon as only half the track is observed (path length 500 m for $P_{1m}(dE;E)$), the resolution becomes similar to using only two energy losses, $P_{500m}(dE;E)$ in Figure 5.18. There is also a bump seen at 10^4 GeV from inconsistencies in the reconstruction distribution which are more noticeable as the path length becomes smaller.

5.3 Idealised Case of Truncated Mean Energy Loss Rate

The truncated energy method (TE) as implemented in IceCube analyses is the robust method described in Section 3.4. In this chapter a simplified version of the TE method is tested against our Edepillim method. This will not give the same resolution results as have been published for the truncated method as the IceCube method was optimised for detector conditions. Later in Chapter 6 Edepillim's resolution will be tested against

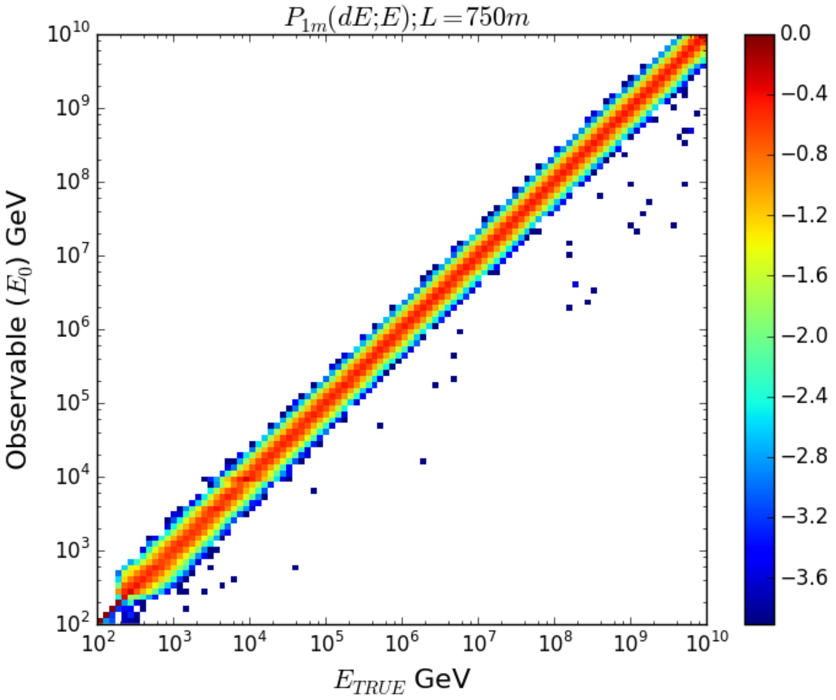


Figure 5.20: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 1\,\mathrm{m}$ and total path length $L = 750\,\mathrm{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}. The distribution still has a small width as with events with path length $L = 1000\,\mathrm{m}$ in Figure 5.14.

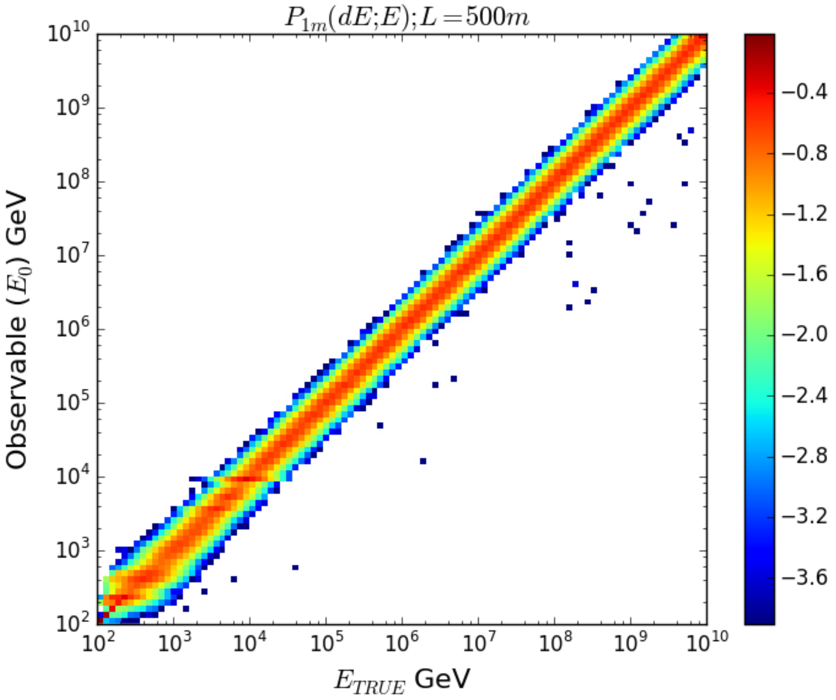


Figure 5.21: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 1\,\mathrm{m}$ and total path length $L = 500\,\mathrm{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}. The distribution has begun to widen with the shortened path length, most noticeably at muon energies less than $10^4\,\mathrm{GeV}$.

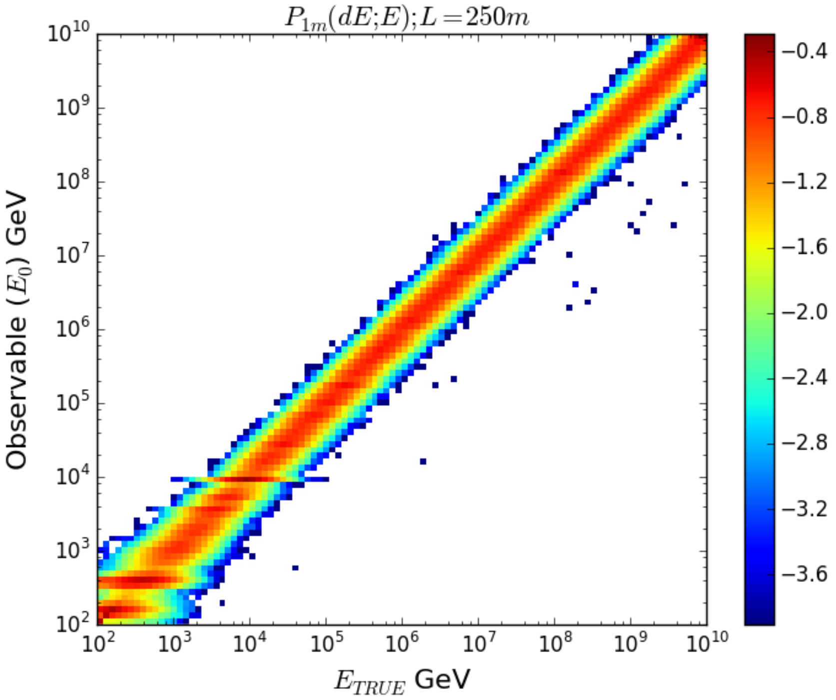


Figure 5.22: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon pattern energy losses of segment length $X = 1\,\mathrm{m}$ and total path length $L = 250\,\mathrm{m}$. Displayed in the logarithm of muon events, normalised per E_{TRUE}. The reconstruction distribution has become wider than using the whole path length in Figure 5.14 with inconsistencies in the reconstruction at energies less than 10^4 GeV

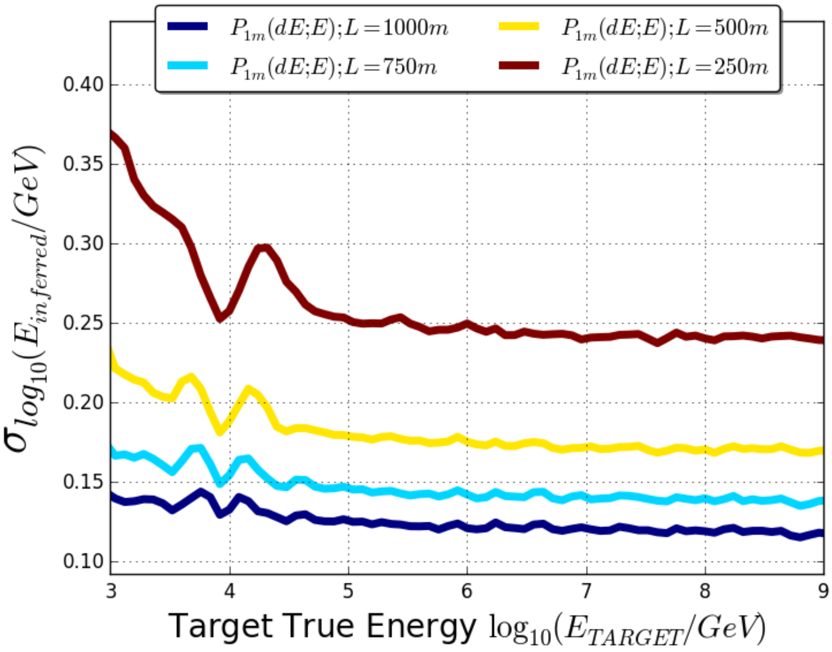


Figure 5.23: The Edepillim energy reconstruction ALF resolution distribution for segment length $X = 1\,\mathrm{m}$ with varying muon path length L over a range of target true muon energies. The bump in the resolution at $10^4\,\mathrm{GeV}$ is caused by inconsistencies in the reconstruction of this energy. This effect is not noticeable at large path lengths but as the path length increases it becomes more obvious.

that from a complete version of the TE method, both methods using the full IceCube simulation and a Millipede reconstruction of the energy loss pattern.

Here we continue with the idealised muon simulation, used in the previous sections, where we use the true energy loss pattern. The TE method operates by removing a number of the largest losses, T, and then calculating the mean energy loss rate dE/dX using the remaining losses. The removal of the losses results in a more uniform energy loss pattern, with a reduction in the number of high energy stochastic losses.

The reconstruction distribution shown in Figure 5.24 is for a TE reconstruction on events of path length $L = 1000\,\mathrm{m}$ and segment length $X = 1\,\mathrm{m}$ while removing the highest loss. Figure 5.24 shows a pattern of underestimating the muon energies, as all true energies are reconstructed at lower energies. While there is a consistent slope of 45 deg, there is also an offset toward lower energies as the number of truncated losses becomes larger. As more losses are removed the overall mean energy loss rate reduces, this is shown in Figure 5.25 where the reconstruction is done on events with path length $L = 1000\,\mathrm{m}$ and segment length $X = 1\,\mathrm{m}$ with the highest 100 losses removed. The offset will not affect the value of the reconstruction resolution as all energies are mapped to the true energy, but it will reduce the range of energies the resolution can be reliably calculated at. This effect is most significant in Figure 5.25 where there is a smaller range of observable muon energies that are reconstructed.

The reconstruction ALF resolution shown in Figure 5.26 demonstrates the improvement in the resolution of the mean energy loss rate from removing one loss $T = 1$ to removing $T = 2$. At $T = 3$ and $T = 4$ there is minimal variation between the resolution results, but as the truncated number increases the resolution results become worse. The optimal number for the truncated loss will vary depending on simulation conditions. For the purpose of testing we will take the optimal TE resolution as $T = 3$.

In Figure 5.27 the effect on energy resolution due to the total path length L is shown, with resolution differences of less than $\Delta\sigma = 0.05$ for path lengths at and above 500 m. At $L = 250\,\mathrm{m}$ there is a larger reduction in the quality of energy resolution. This shows that like the Edepillim method the TE method's resolution is also dependent on the path length.

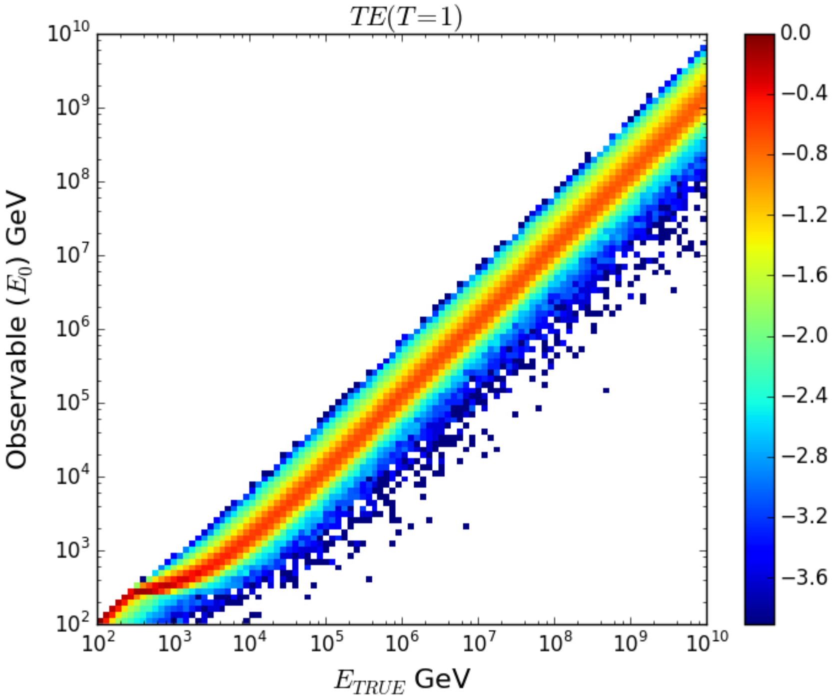


Figure 5.24: Reconstruction distributions for an idealised TE of the observable muon reconstructed energy E_0 vs E_{TRUE}. Displayed in the logarithm of muon events, normalised per E_{TRUE}. E_0 is calculated as the mean energy loss rate (dE/dX), with $T = 1$ of the largest energy losses truncated from the muon pattern energy losses, with segment length $X = 1\,\mathrm{m}$ and path length $L = 1000\,\mathrm{m}$.

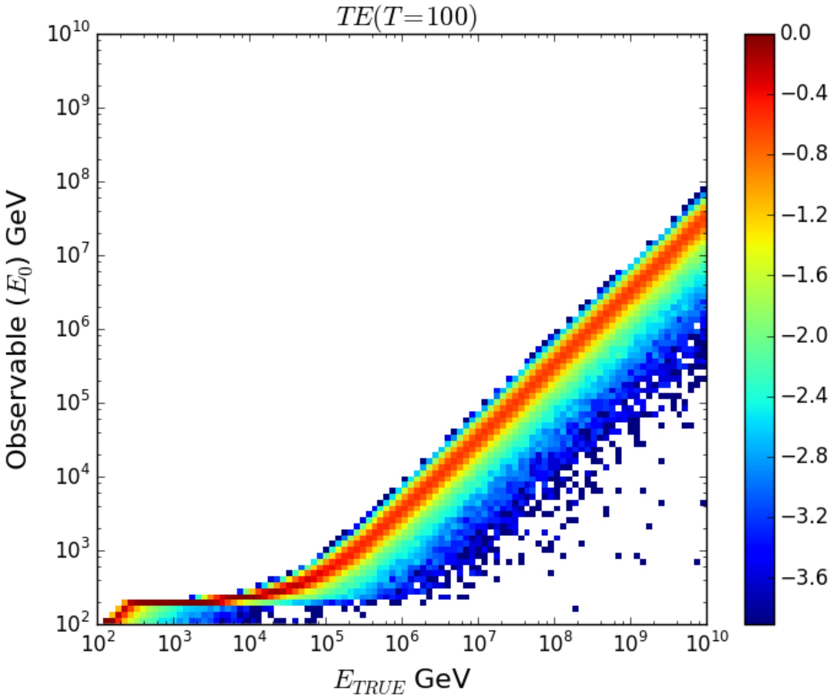


Figure 5.25: Reconstruction distributions for an idealised TE of the observable muon reconstructed energy E_0 vs E_{TRUE}. Displayed in the logarithm of muon events, normalised per E_{TRUE}. E_0 is calculated as the mean energy loss rate (dE/dX), with $T = 100$ of the largest energy losses truncated from the muon pattern energy losses, with segment length $X = 1\,\mathrm{m}$ and path length $L = 1000\,\mathrm{m}$. At this truncation there is an offset in the reconstructed energy with respect to true energy. This is an extreme case of truncation and would not be optimal truncation level.

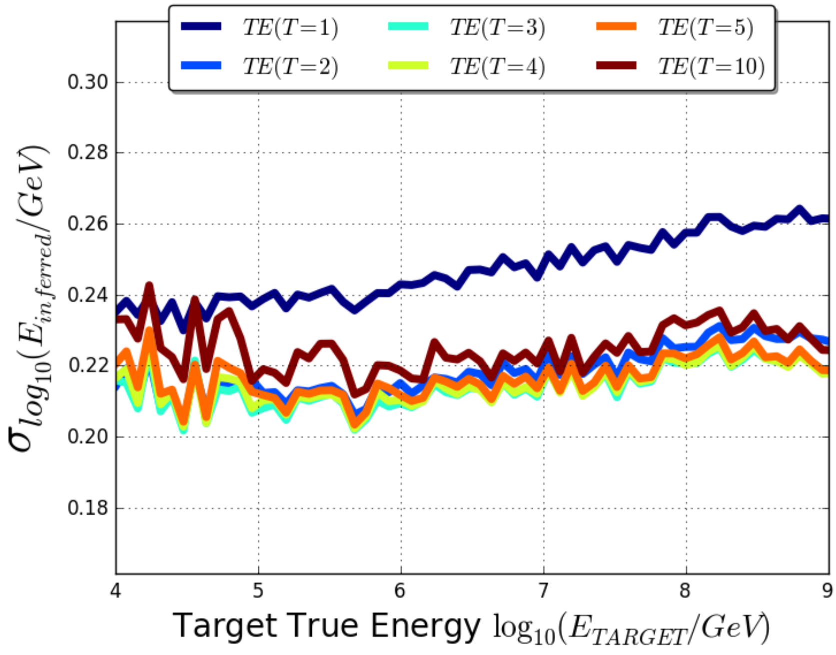


Figure 5.26: The TE ALF reconstruction resolution for target true energies for events with path length $X = 1000\,\text{m}$ and segment length $1\,\text{m}$ for the cases of truncating $T = 1, 2, 3, 4, 5, 10$ losses. Not shown is $T = 100$ as the resolution is above the range of the others, with $\sigma = 0.25 - 0.35$. The optimal truncation value is $T = 3$ for this simulation as it has the best reconstruction resolution.

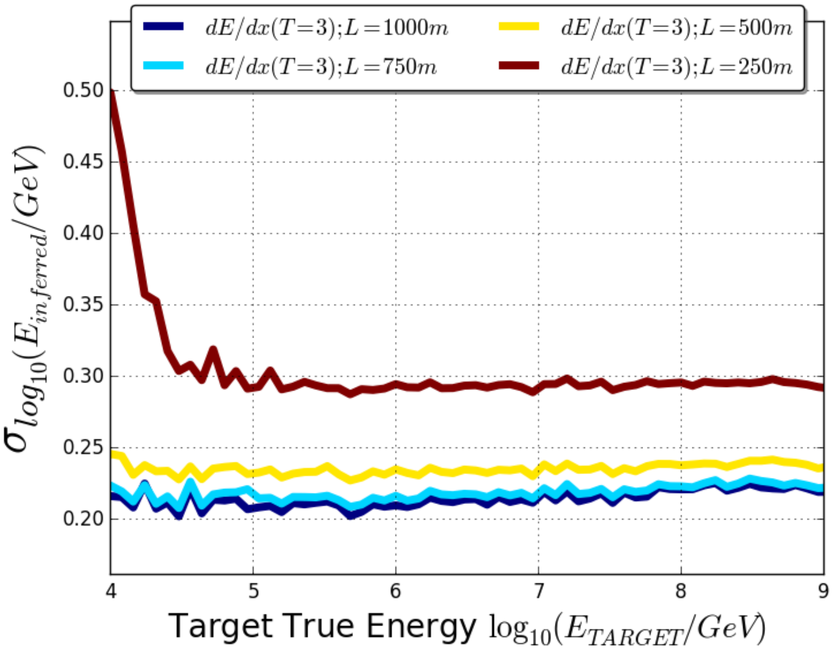


Figure 5.27: The TE ALF reconstruction resolution for target true energies for events with path length $L = 1000, 750, 500, 250\,\mathrm{m}$ and segment length $X = 1\,\mathrm{m}$ for the case of truncating $T = 3$ losses. There is only small changes in the reconstruction resolution for $L = 1000, 750, 500\,\mathrm{m}$ but at $L = 250\,\mathrm{m}$ the resolution is much worse. TE thus follows the same pattern as the Edepillim method with a worsening resolution with decreased path length.

5.4 Comparison of Results

The results of the energy resolution can be compared for the two energy reconstruction methods, Edepillim and TE. The resolutions in Figure 5.28 are for the optimal segment length or truncated loss that was found for each method. For Edepillim the optimum is segment length $X = 1\,\text{m}$, and TE removes the three highest losses ($T = 3$).

In the idealised case Edepillim has the best resolution of the two methods, however this resolution is limited by the level of bin size resolution in the energy loss pattern, if the losses were known to a smaller segment length than $X = 1\,\text{m}$ the reconstruction resolution would improve. While truncated energy with the three highest losses removed performs with a slightly larger resolution, it has been proven to be a robust and consistent method to perform within IceCube [34]. It is also important to note that the difference between Edepillim and TE is only $\sigma_{TE} - \sigma_{Edepillim} \sim 0.1$ which means all real-life effects of the detector response must be considered, in order to decide on the best reconstruction in practice. This study is the subject of the next chapter.

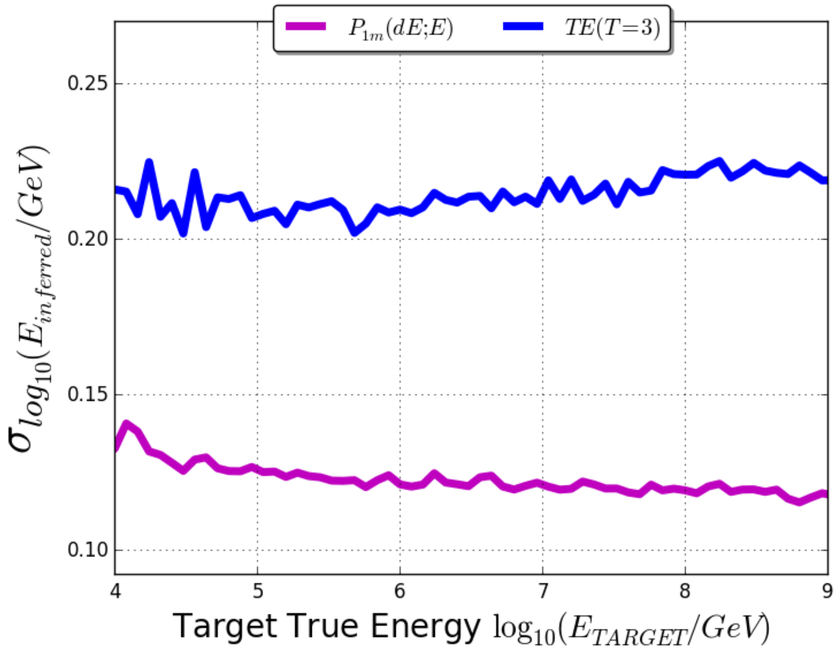


Figure 5.28: The ALF reconstruction resolution for comparison of the Edepillim method against the optimal TE method for muon energy pattern path length $L = 1000\,\text{m}$.

5.5 Summary

The Edepillim method uses a maximum likelihood approach to reconstruct the muon energy at the beginning of an energy loss pattern. The method relies on accurate descriptions of the muon energy losses along the path length using probability distribution functions built from simulation to describe the energy losses given the muon energy. Using the energy loss pattern, Edepillim can perform a maximum likelihood fit over a range of test energies to find the reconstructed muon energy. This method produces a very good reconstruction resolution with tests having been made for various segment lengths showing that the smallest segment length will give the best result. Additionally, the total muon path length was tested showing that the longest path length gave the best result. The Edepillim reconstruction will continue to improve with more information about the true energy losses. That is the more individual energy losses the better the reconstruction, with improvements seen for longer path lengths and smaller segment length rebinning.

The TE reconstruction is an alternative method that was tested against the Edepillim method for this idealised case. The TE method in this idealised case does not perform as well as the Edepillim method. The Edepillim method performs with the best resolution across a range of energies with the greatest potential for improved resolution with more information on the energy losses.

Chapter 6

Edepillim Energy Reconstruction on the IceCube Simulation

The Edepillim muon energy reconstruction method as described in Chapter 5 is a topological method that uses all the muon energy losses to reconstruct the initial energy, resulting in a good reconstruction resolution. Such a method would be ideal for the IceCube neutrino detector, as the large detector size means that for each muon event there is potentially a large energy loss pattern that can be reconstructed for that track. As such Edepillim is applied to a simulation of events detected using IceCube, to explore its effectiveness.

The initial testing of the Edepillim reconstruction on the IceCube simulation resulted in a much worse resolution than for the truncated energy method (TE). This was unexpected as the method performed with a good resolution on simulations where the true energy losses were known, described in Chapter 5. However when performed on the IceCube simulation many more effects are needed to be taken into account. The energy losses are not true losses but reconstructed via the Millipede method described in Section 3.5. This introduces uncertainty in the accuracy of the energy losses, both in size and position of the energy loss, and these affect the performance of Edepillim. We also highlight the impact that an energy loss pattern with a majority of zero energy losses has on the reconstructed energy.

In this Chapter we describe the methods explored in order to get Edepillim to function with a better resolution on both the true losses and reconstructed losses. For true losses this includes accounting for the continuous energy losses along the energy loss pattern such that a consistent PDF can be implemented.

For the reconstructed losses the position of an energy loss within the volume of the detector is taken into account to remove inaccurate energy losses. The method that has the most improvement on the reconstruction is the rebinning of the energy loss segment lengths to create a more accurate energy loss pattern. As a final improvement to the resolution, energy losses that are at the threshold of the PDF are not included in the likelihood.

Finally two selection criteria for the events are explored, one involving selecting events based on the Edepillim likelihood space and the other selection based on the path length of the muon events. The selection of events using path length was found to have the most improvement in reconstruction resolution. The final results demonstrate Edepillim working with a better resolution than TE for a selection of muon events.

The Chapter's final section outlines an alternative approach of using the Millipede reconstructed energy losses to build the PDF, such that the errors of the energy losses will be included in the PDF. This gives the same reconstruction resolution to that of other sections that describe methods of improving the estimate of the reconstructed energy losses, because the error in the reconstructed losses has now been accounted for in the previous sections such that the reconstruction works with the PDF built from Millipede losses.

6.1 Edepillim Initial Results

In this chapter we test the Edepillim method using a simulation of events in which muons starting outside the IceCube detector will travel through the detector in all directions. Each event has its energy loss pattern reconstructed using Millipede with SplineMPE used as the muon track direction method, as outlined in Section 3.5.

The Edepillim method works by reconstructing the muon energy from the energy losses that are reconstructed using Millipede. The method follows the same implementation laid out in Chapter 5 with the following variations. Since these are reconstructed losses it cannot be assumed that energy losses of zero at the end of the energy loss pattern indicate a muon that has lost all of its energy, so no such assumption will be included in the reconstruction. The PDFs used in the maximum likelihood scan will be smoothed using the interpolation fitter outlined in Appendix C.

For all comparisons the truncated mean energy loss reconstruction (TE), on the same events is shown. This energy reconstruction is the standard method used in IceCube, [34], and the optimal settings for TE used in this work are described in Appendix B.

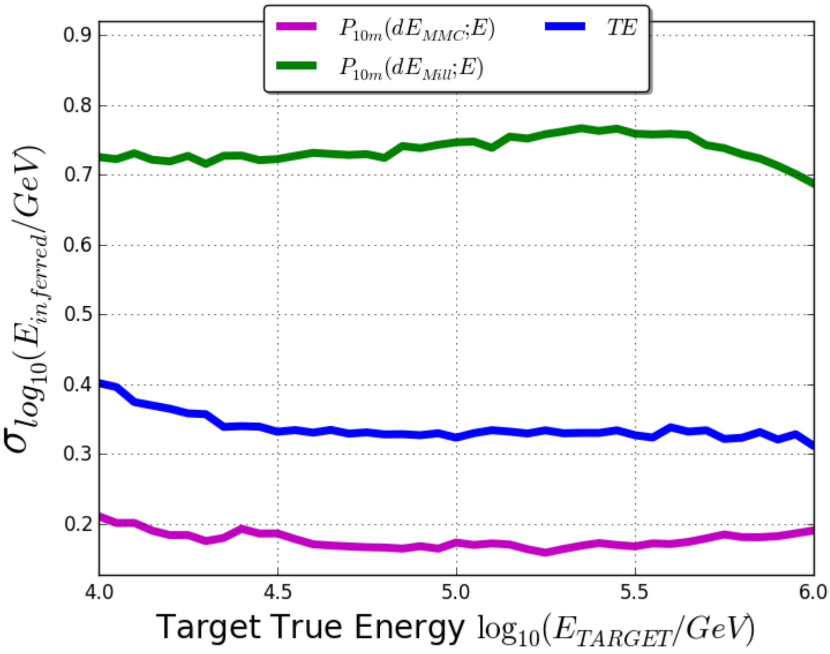


Figure 6.1: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the true losses ($d\vec{E}_{MMC}$) and the reconstructed losses ($d\vec{E}_{Mill}$). The standard IceCube method (TE) is shown for comparison. The resolution for the Edepillim reconstruction on Millipede losses are worse than the TE, and the resolution for the true losses is at a higher value of σ than expected based on the results in Chapter 5 which were also done on idealised muons where the true losses are known.

The reconstruction resolution shown in Figure 6.1 is not what was expected based on the results from Chapter 5, where the Edepillim reconstruction on true losses ($P_{10m}(dE_{MMC}; E)$) was performing with better resolution than TE in the idealised case. However the reconstruction distribution for the true energy losses in Figure 6.2 shows a wider distribution than it should, especially at energies below 10^4 GeV; the distribution should be similar to that of Figure 5.15 which is also a reconstruction distribution on idealised true losses for the same segment length.

Additionally the energy reconstruction on the Millipede reconstructed losses ($P_{10m}(dE_{Mill}; E)$) has an extremely poor value of resolution σ compared to that using the true losses. The reconstruction distribution in Figure 6.3 shows a large width of the distribution with a large scattering of low energy muons being reconstructed to high energy. Again this is not what was expected based on the simulation studies in Chapter 5 and indicates a problem in the energy loss pattern that is causing the loss in reconstruction resolution.

The initial results show that the Edepillim reconstruction on Millipede reconstructed energy losses has a very poor reconstruction resolution compared with using TE. Also the Edepillim reconstruction on true energy losses is giving a worse resolution than expected using true losses, as was shown in Chapter 5, where it produced a good energy resolution. The cause of this loss in resolution will be explored in this chapter, with descriptions of solutions that were implemented to improve the Edepillim resolution. The notation representing each improvement to the method given in this chapter can be found in Table 6.1 with brief descriptions.

Reconstruction Notation	
Reconstruction	Description
$P_{10m}(dE_{(pattern)}; E)$	Edepillim reconstruction on an energy loss pattern $dE_{pattern}$ where *pattern* is either; Millipede reconstructed energy losses (dE_{Mill}) or the true energy losses (dE_{MMC}), for segment length 10 m. Using a PDF built from the stochastic energy losses in PROPOSAL of segment length 10 m.
$P_{10m}(V_{pattern}; E)$	Edepillim reconstruction on an energy loss pattern for segment length 10 m. Using a PDF built from the stochastic energy losses in PROPOSAL of segment length 10 m binned in fractional energy $V = dE/E$.

Reconstruction Notation continued	
Reconstruction	Description
$P_{10m}(a)(dE_{pattern}; E)$	Edepillim reconstruction on Millipede/true energy losses accounting for continuous energy losses. Using a PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length 10 m
$P_{10m}(b)(dE_{pattern}; E)$	Edepillim reconstruction on Millipede/true energy losses accounting for continuous energy losses, with only losses contained within the detector volume in the energy loss pattern. Using a PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length 10 m.
$P_{10m}(dE_{MMC}(\%); E)$	Edepillim reconstruction on true energy losses accounting for continuous energy, with only a percentage of highest energy true losses, the remaining are threshold energy losses, $dE_{threshold} = 2\,\text{GeV}$. Using a PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length 10 m.
$\sum d\vec{E}_{pattern}$	Reconstructed muon energy is given by the sum of all the energy losses in the pattern, where *pattern* is Millipede($Mill$) or true losses (MMC).
$P_X(b)(dE_{pattern}; E)$	Edepillim reconstruction on Millipede/true energy losses, accounting for continuous losses and contained within the detector volume. For segment length > 10 m the energy loss pattern has been rebinned to create larger segment length energy loss patterns. Using a corresponding PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length X.

Reconstruction Notation continued

Reconstruction	Description
$P_X(c)(dE_{pattern}; E)$	Edepillim reconstruction on Millipede/true energy losses, accounting for continuous losses and contained within the detector volume. With losses below the threshold, $dE_{threshold} < 0.2 \times X$, removed from the likelihood calculation. For segment length $> 10\,\mathrm{m}$ the energy loss pattern has been rebinned to create larger segment length energy loss patterns. Using a corresponding PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length X.
$P_X(d)(dE_{pattern}; E)$	Edepillim reconstruction on Millipede/true energy losses, accounting for continuous losses and contained within the detector volume, with only events which pass the likelihood selection criteria, $E_0 > 2 \times \sum d\vec{E}_{Mill}$. For segment length $> 10\,\mathrm{m}$ the energy loss pattern has been rebinned to create larger segment length energy loss patterns. Using a corresponding PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length X.
$P_X(c)(dE_{pattern}; E); L$	Edepillim reconstruction on Millipede/true energy losses, accounting for continuous losses and contained within the detector volume. With losses below the threshold, $dE_{threshold} < 0.2 \times X$, removed from the likelihood calculation. With a selection cut done on the path length L. For segment length $> 10\,\mathrm{m}$ the energy loss pattern has been rebinned to create larger segment length energy loss patterns. Using a corresponding PDF built from continuous energy losses in PROPOSAL, $E_{initial} - E_{final}$, of segment length X.

Reconstruction Notation continued	
Reconstruction	Description
$P_X(Mill)(dE_{pattern}; E); L$	Edepillim reconstruction on Millipede/true energy losses, accounting for continuous losses and contained within the detector volume. For segment length $> 10\,\text{m}$ the energy loss pattern has been rebinned to create larger segment length energy loss patterns. Using a corresponding PDF built from the reconstructed Millipede energy losses of segment length X.

Table 6.1: Table of the labels and descriptions of the variations of reconstructions used in this chapter.

6.2 Millipede Reconstructed Muon Energy Losses

The Edepillim energy reconstruction relies on PDFs that are based on the true energy losses expected for muon interactions, thus if the reconstructed losses do not accurately describe these it will affect the overall performance of the Edepillim reconstruction. Therefore an understanding of how the muon event's energy loss pattern has been reconstructed is essential. IceCube uses the Millipede software as described in Section 3.5 to reconstruct the energy loss pattern assuming the event direction from the SplineMPE fit. Millipede uses a regularised unfolding on all the waveforms in the DOMs to reconstruct the energy losses at every point along the track, the only constraint applied is that the losses have to be non-negative.

The overall resolution of Edepillim reconstructions on Millipede events is much worse than both the TE and Edepillim on the true energy losses, as shown in Figure 6.1. The cause of this difference of resolution between the two Edepillim reconstructions most likely comes from the difference between the Millipede energy loss pattern ($d\vec{E}_{Mill}$) and the true energy loss pattern ($d\vec{E}_{MMC}$).

There are several trends in the Millipede reconstructed losses that can be observed in the reconstructed energy loss patterns shown in Figures 6.4 to 6.8. The most obvious is shown in Figure 6.4 which represents a "standard" energy loss pattern that would be

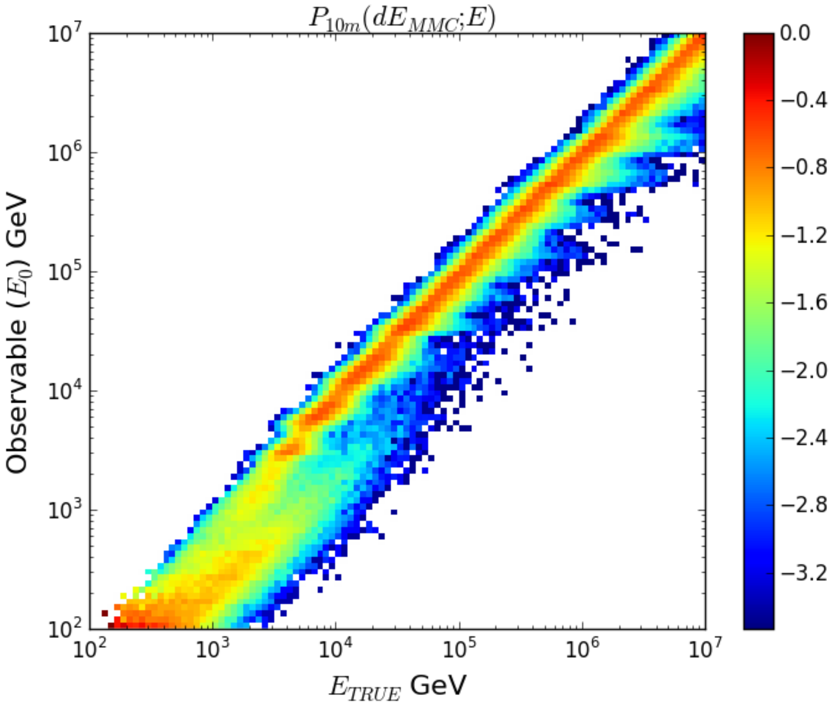


Figure 6.2: Edepillim reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE}, when using the true energy loss pattern (MMC), displayed in the logarithm of muon events normalised along E_{TRUE}. For this reconstruction it is expected that this distribution would have a similar width in E_0 and E_{TRUE} to that of the idealised case in Figure 5.15, however this distribution is wider.

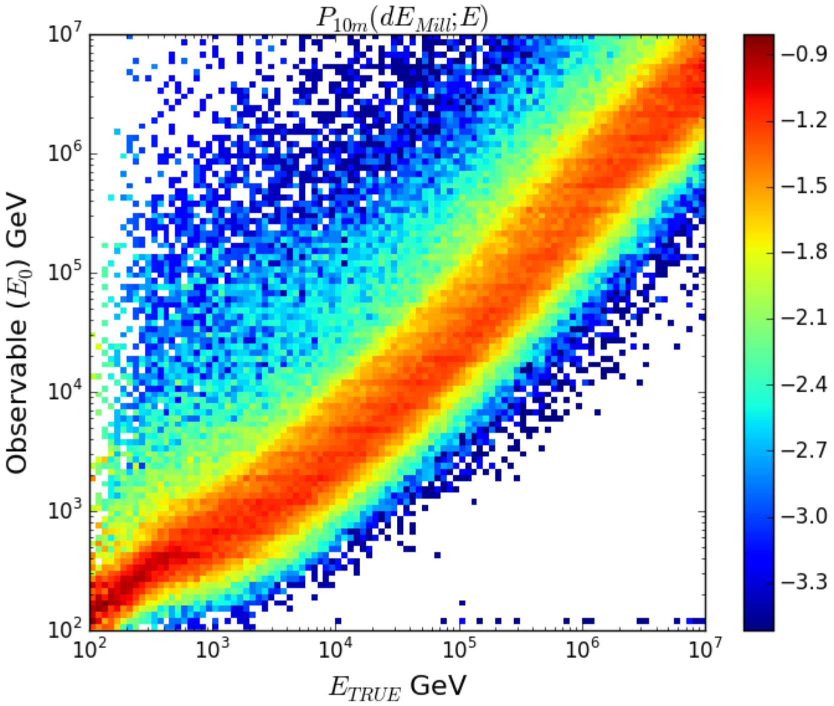


Figure 6.3: Edepillim reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE}, when using the reconstructed energy loss pattern (Mill). Displayed in the logarithm of muon events normalised along E_{TRUE}. There is a large amount of spread in the distribution, with some events having a reconstructed energy larger than their true initial energy, causing the spread towards the top left corner.

reconstructed. It shows the Millipede (*green*) tendency to favour discrete larger losses with surrounding bins containing either a zero energy loss or a loss orders of magnitude smaller. This is a general trend that can be seen in all other energy loss patterns. In the true case the losses (*pink*) tend to be more continuous despite containing only stochastic energy losses. Figure 6.4 also shows the detector boundary (dashed lines), beyond these bounds the energy losses occur outside the IceCube detector but the scattered photons from these losses can be detected by DOMs on the edge of the detector. This allows for a longer energy loss pattern than what is actually contained within the detector to be reconstructed.

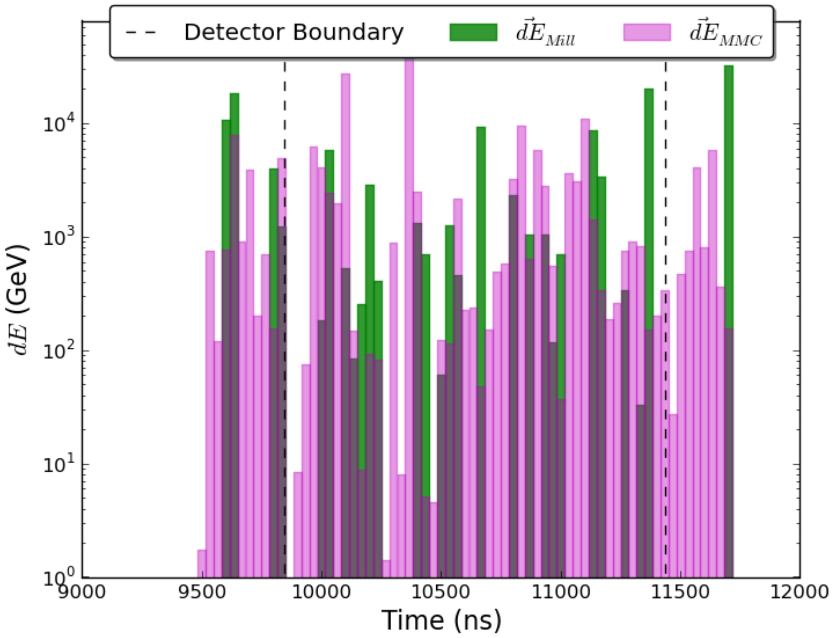


Figure 6.4: A simulated muon's energy loss pattern showing its true losses $d\vec{E}_{MMC}$ (*pink*) and its reconstructed losses $d\vec{E}_{Mill}$ (*green*), for an event with energy $E_{TRUE} = 1.2 \times 10^6\,\text{GeV}$ and path length $L = 680\,\text{m}$, based on the dashed lines which show the detector boundary for the track. This event has reconstructed losses that approximate the amplitude of the true losses, with offsets to the position of the losses along the energy loss pattern. There are also a number of Millipede reconstructed losses that are zero in the pattern.

In Figure 6.5 a low muon energy event of initial energy $E_{TRUE} = 7.3 \times 10^3\,\text{GeV}$ is shown. In this event even fewer Millipede losses (*green*) are seen to be correctly reconstructed, and the larger Millipede losses are often reconstructed in different time bins than in the true loss case (*pink*).

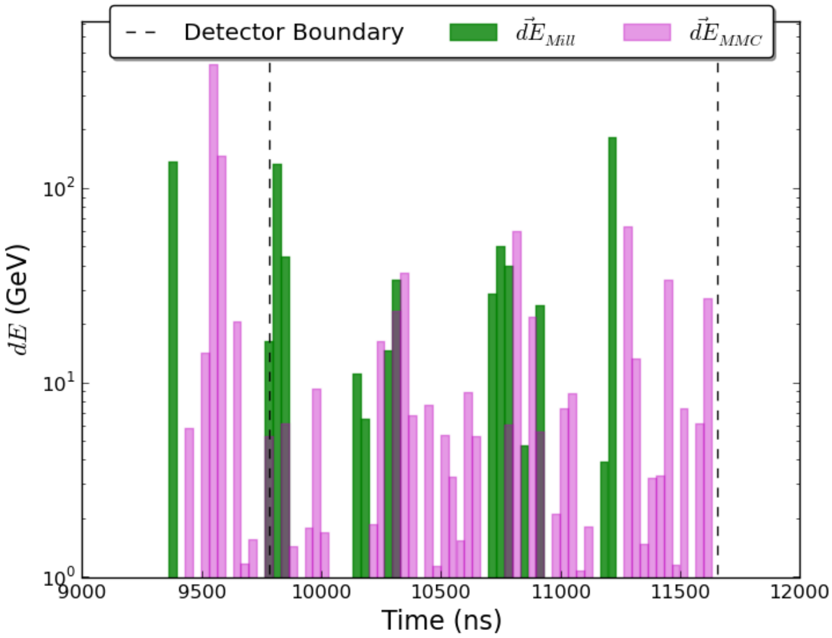


Figure 6.5: A simulated muon's energy loss pattern showing its true losses $d\vec{E}_{MMC}$ (*pink*) and its reconstructed losses $d\vec{E}_{Mill}$ (*green*), for an event with energy $E_{TRUE} = 7.3 \times 10^3\,\mathrm{GeV}$ and path length $L = 700\,\mathrm{m}$. The dashed lines show the detector boundary for the track. The majority of the reconstructed losses $d\vec{E}_{Mill}$ are zero.

Another feature that hampers the Edepillim method is shown in the case of Figures 6.6 and 6.7 where Millipede reconstructs a large energy loss, several orders of magnitude greater than any loss in the true energy loss pattern. This is seen in Figure 6.6 with a $dE \sim 10^9\,\mathrm{GeV}$ loss at the end of the pattern, and in Figure 6.7 as a $dE > 10^6\,\mathrm{GeV}$ loss at the beginning of the pattern. Occasionally Millipede interprets small detections of photoelectrons in the waveforms, usually from DOMs at the edge of the detector, as being caused by a large stochastic loss outside or at the edge of the detector. In the further example given in Figure 6.8 this large loss can also be placed inside the detector boundary despite no large true loss in or near that time bin.

In the case of Figure 6.7 this particular example produces a large loss outside the detector that is 4 orders of magnitude greater than the next loss. Due to Edepillim's physical constraint, of always reconstructing to an energy greater than the sum of the energy losses, it would not allow for reconstruction lower than this loss, therefore the starting energy of the likelihood scan would be at minimum 4 orders of magnitude greater than the sum of the true energy losses. This will cause some events to be well out of range of their true energy, such that the true energy of the event may not be in the likelihood scan. This explains much of the large spread towards higher reconstructed energies than the true muons energies, seen in the top left of Figure 6.3.

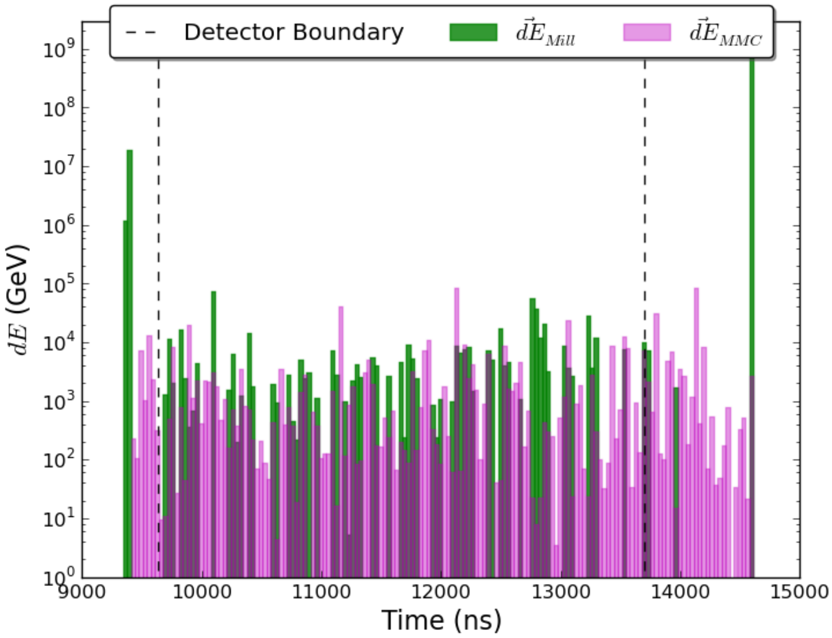


Figure 6.6: A simulated muon's energy loss pattern showing its true losses $d\vec{E}_{MMC}(pink)$ and its reconstructed losses $d\vec{E}_{Mill}$ $(green)$, for an event with energy $E_{TRUE} = 1.7 \times 10^6$ GeV and path length $L = 1630$ m. The dashed lines show the detector boundary for the track. This event has three large energy losses that occur outside the detector volume which are orders of magnitude larger than the true losses occurring in the same spot.

The following sections describe how the features of the Millipede reconstructed energy loss pattern create problems with the Edepillim energy reconstruction. Also explored are the methods that can be used in order to allow Edepillim to perform with improved resolution.

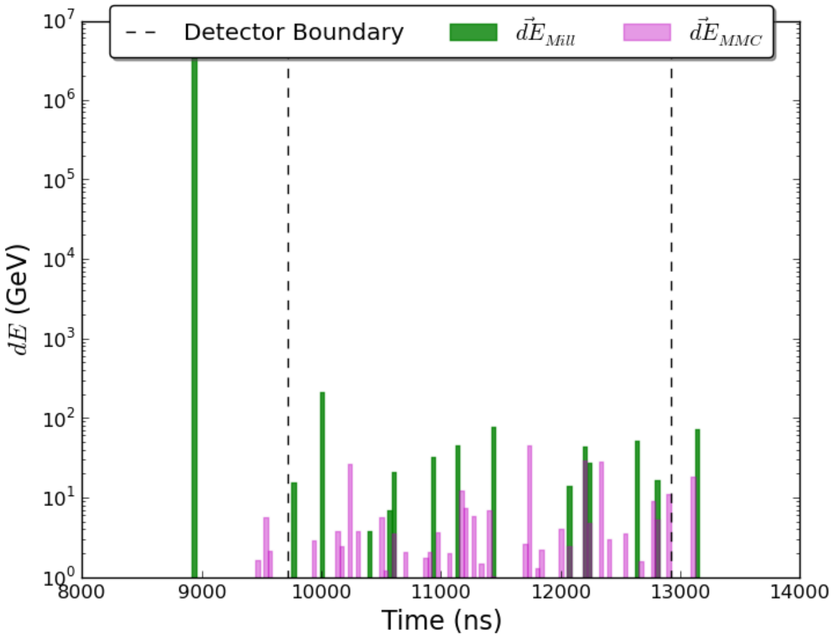


Figure 6.7: A simulated muon's energy loss pattern showing its true losses $d\vec{E}_{MMC}$ (*pink*) and its reconstructed losses $d\vec{E}_{Mill}$ (*green*), for an event with energy $E_{TRUE} = 2.2 \times 10^3$ GeV and path length $L = 1270$ m. The dashed lines show the detector boundary for the track. Like the event shown in Figure 6.5 this event has a low number of reconstructed losses, also there is a large energy loss reconstructed outside the detector boundary, seen at just before 9000 ns and this loss is five orders of magnitude greater than any true energy loss.

6.3 The Zero Energy Loss Problem

As can be seen in the reconstruction distribution in Figure 6.2 based on the true energy losses recorded in the simulation, it is not just the reconstructed losses causing problems for Edepillim. The true losses should have produced a resolution closer to the resolution in Chapter 5, which is an idealised case using true losses in Figure 5.15, but the reconstruction resolution in Figure 6.2 does not produce the same level of resolution. Because it does not it indicates a problem with the energy reconstruction on the energy loss pattern for the true losses, particular for the low energy muons below 10^4 GeV. Here we discuss how the loss in resolution comes from how the zero energy loss bins are treated.

The range of true energy losses for the entire simulation of muon events can be seen

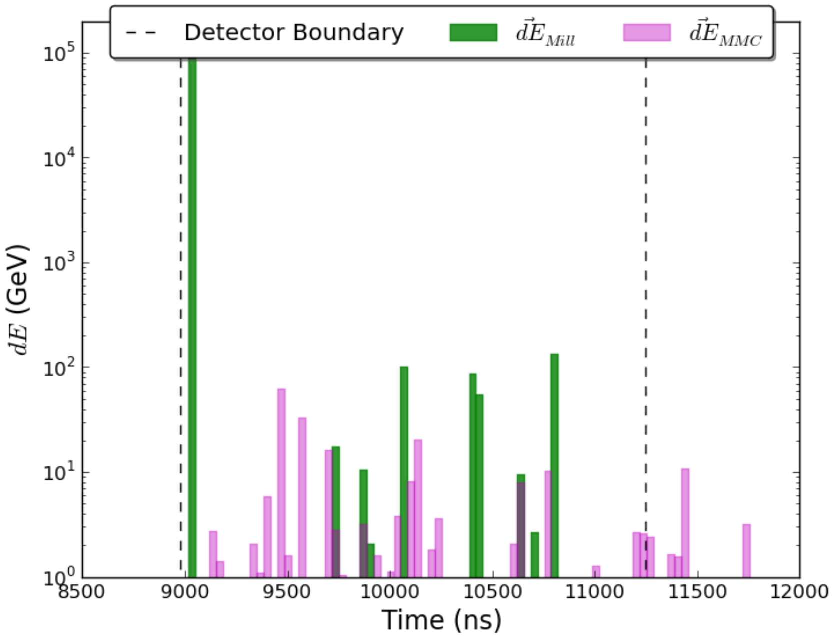


Figure 6.8: A simulated muon's energy loss pattern showing its true losses $d\vec{E}_{MMC}$ (*pink*) and its reconstructed losses $d\vec{E}_{Mill}$ (*green*), for an event with energy $E_{TRUE} = 1.1 \times 10^3$ GeV and path length $L = 850$ m. The dashed lines show the detector boundary for the track. This event has a large energy loss three orders of magnitude greater than any of the true losses. Unlike Figure 6.6 and 6.7 this loss, near 9000 ns, has been reconstructed inside the detector boundary.

in Figure 6.9 where there is a break in the range of losses between approximately 0.4×10^{-2} GeV to 0.4×10^{-1} GeV. The losses in this range have not been recorded in the simulation and thus cannot be reproduced in the energy loss pattern. These missing losses can potentially leave gaps in an energy loss pattern. Not shown in Figure 6.9 are the contribution of all the zero energy loss bins which are 23% of the total energy losses. The break in the histogram if it was continuous would make up approximately 2.4% of the total energy loss bins, thus a small portion of the energy losses were not recorded and when the losses are binned into an energy loss pattern it will result in zero loss bins. But the more significant factor to the reconstruction is that there are approximately 20% of its total energy loss bins that contain zero energy losses due to no stochastic losses occurring in that bin.

The energy loss patterns are not made as they were in Chapter 5, where PROPOSAL returned the full energy loss in a segment length. Instead every individual stochastic energy loss that occurs along the muon length is recorded with its type, energy, position and time in a data tree of the primary and daughter particles of the event, known as an I3MCTree. In order to produce an energy loss pattern for this analysis the losses in the tree are sorted by time and each loss in the same time bin as the Millipede energy loss pattern will be combined. The problem with the true energy loss pattern is that not every loss for the daughter particles will be saved in the I3MCTree, including any continuous losses such as ionisation. It is because of the large size of the muon event tree that it is considered unneeded to recorded every energy loss. However when the true pattern of energy loss is produced for this analysis the lack of every loss will cause some bins in the pattern to have zero energy loss or a loss smaller than the bounds of the PDF. This is mostly seen to cause the effect in Figure 6.2 with the large width of the energy reconstruction distribution.

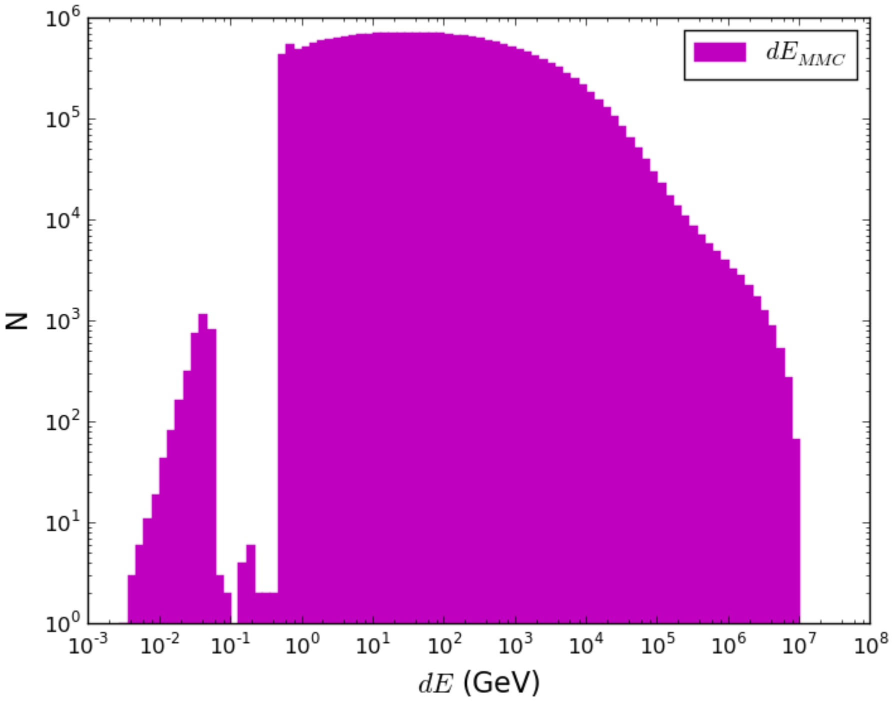


Figure 6.9: Histogram of all the individual true energy losses (dE) in the simulation set. No distinction has been made between the different muon events. Not displayed in the above log plot are the zero energy losses which make up 23% of the total energy loss bins.

As mentioned in Section 6.2 there is a tendency in the Millipede reconstruction energy loss pattern to have a majority of zero energy loss bins. It also has a tendency to reconstruct extremely small losses as seen in Figure 6.10. These losses have too small a range to be considered physical. Again not shown in Figure 6.10 are the contributions of zero energy losses. Reconstructed energy losses have an abundance of zero energy losses, with them being 71% of all the energy losses.

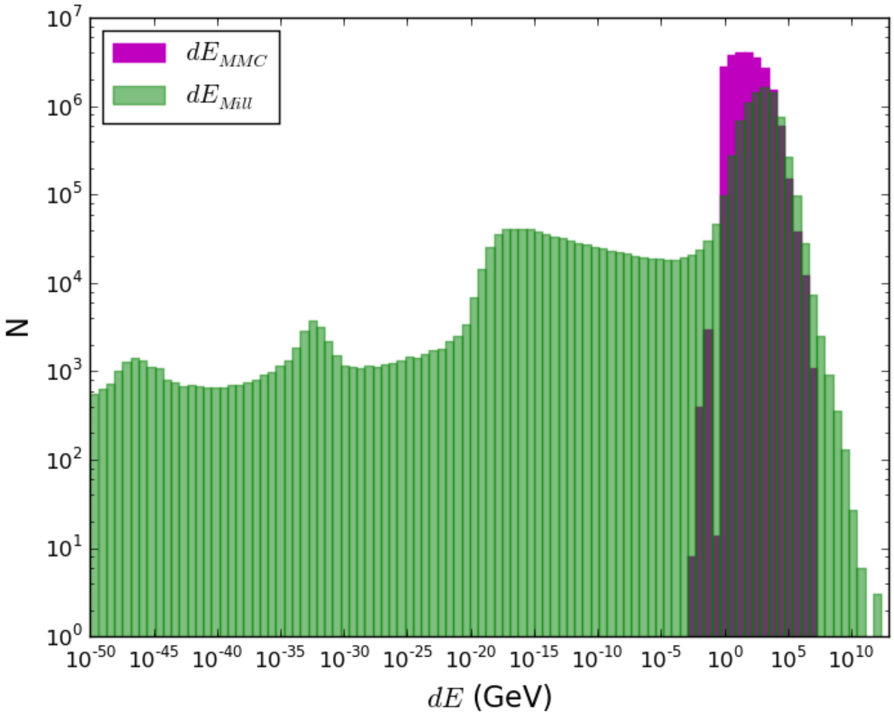


Figure 6.10: Histogram of all the individual energy losses (dE) in every muon event in the simulation set, for true losses (dE_{MMC}) and reconstructed energy losses (dE_{Mill}) for segment length $X = 10\,\text{m}$. No distinction has been made between the different muon events. There are a number of losses that are below 10^{-3} GeV which do not correspond to any true energy losses. Not displayed in the above log plot are the zero reconstructed energy losses which make up 71% of the total Millipede reconstructed energy loss bins.

All of these zero and small energy losses cause problems with the Edepillim reconstruction and PDFs. In this section we discuss the technical implications of these losses. In Section 6.5 the reconstruction application implications of such losses are discussed.

The problem in the technical application of these losses is that the PDFs are built from PROPOSAL simulation (see Section 3.6). Unlike in Chapter 5, the PDFs here ignored continuous energy losses. In Chapter 5 these were included by using $dE = E_{inital} - E_{final}$ for the loss in each segment length. Thus in the initial testing in Section 6.1 the energy loss entered into the PDF is the total of the stochastic losses in a given segment length X from PROPOSAL, but does not contain the continuous losses you would get from using $dE = E_{initial} - E_{final}$ from PROPOSAL. Despite using only stochastic losses there is still a limit to the smallest loss that PROPOSAL returns for the low energy muons as seen in Figure 6.11. At low muon energies below $10^3\,\mathrm{GeV}$ there is always a minimum loss that PROPOSAL predicts of the order of $10^{-1}\,\mathrm{GeV}$ even in the stochastic case. At higher muon energies above $10^4\,\mathrm{GeV}$ there are cases where the losses fall into the underflow bin of the PDF; all these losses are the rare case where there is no stochastic loss in the small segment length $X = 10\,\mathrm{m}$. This does not occur at higher segment length as it is more likely for a stochastic loss to occur.

Due to the construction of the PDFs being between an energy loss range $10^{-1} < dE < 10^{10}$, when there is a reconstructed zero energy loss, $dE = 0$, there is no corresponding probability bin in the PDF. The nearest thing to this probability is to make the assumption that it can be approximated by the probability of the smallest energy loss in the range, which for the PDF in Figure 6.11 is $dE_{min} = 10^{-1}\,\mathrm{GeV}$. While the fitted nature of the PDFs (see Appendix C) means a probability will be given for all losses, it does not mean it will be the correct probability. A fitted PDF will tend to smaller numbers at the edge of the spline if there are no data points, meaning that for a muon event of the probability $P(0; E = 10^2\,\mathrm{GeV})$ (which for a zero loss will be taken as the probability for $dE = 10^{-1}\,\mathrm{GeV}$) will be less than $P(0; 10^4\,\mathrm{GeV})$. This is the cause of the smearing in reconstruction resolution in Figure 6.2 at the lower energy where events will not have an accurate probability for zero losses.

The jump in probability for the zero energy losses becomes more apparent in the case of using the fractional energy loss PDF. In Figure 6.12, due to limitations in the PROPOSAL for producing energy losses at lower energies there is a diagonal line representing the smallest loss that is possible for PROPOSAL, i.e. $dE_{threshold} = 10^{-1}\,\mathrm{GeV}$. The diagonal line then represents a mapping of the threshold energy to the fractional energy bin. In the case of small energy losses, for an interpolated fractional PDF the zero loss will correspond to a very small probability, thus to explain the loss Edepillim will find the

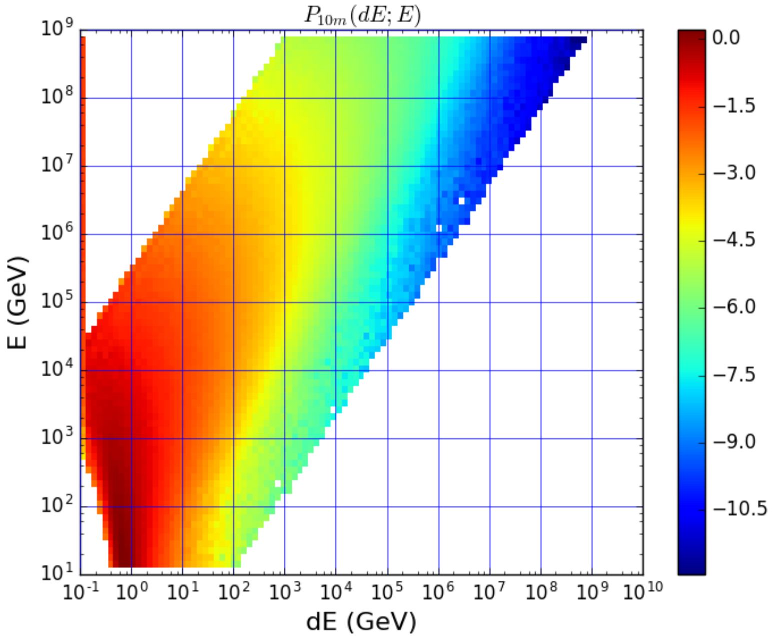


Figure 6.11: Probability distribution function built from PROPOSAL simulated losses of segment length $X = 10\,\mathrm{m}$ binned along energy loss dE. The vertical line at 10^{-1} is an underflow bin of every loss that is beyond the range of the histogram, with losses in this segment range going as low as $dE = 10^{-3}\,\mathrm{GeV}$. The minimum cut on V by PROPOSAL (see section 3.6) causes the diagonal cut off line from muon energies $10^1 - 10^3\,\mathrm{GeV}$, this line would continue to higher muon energies if the range of the PDF was larger. At higher muon energies, above $10^4\,\mathrm{GeV}$, there is a gap between the PDF and the underflow bin, at higher energies stochastic losses that occur are more likely to be larger, which causes the diagonal line above $10^4\,\mathrm{GeV}$. However due to the small segment length there are cases where the are very small stochastic losses and these fall into the underflow bin.

most likely initial energy to be at an energy where a zero loss (which will correspond to the smallest possible energy loss dE_{min}) does have a significant probability i.e above $E = 3 \times 10^3$ GeV. As such Edepillim will have poor reconstruction resolution on any low energy muon events ($E_{TRUE} < 3 \times 10^3$ GeV). This results in some strange behaviour in the reconstruction distribution seen in Figure 6.13, with a clear line at the reconstructed energy $E_0 = 3 \times 10^3$ GeV.

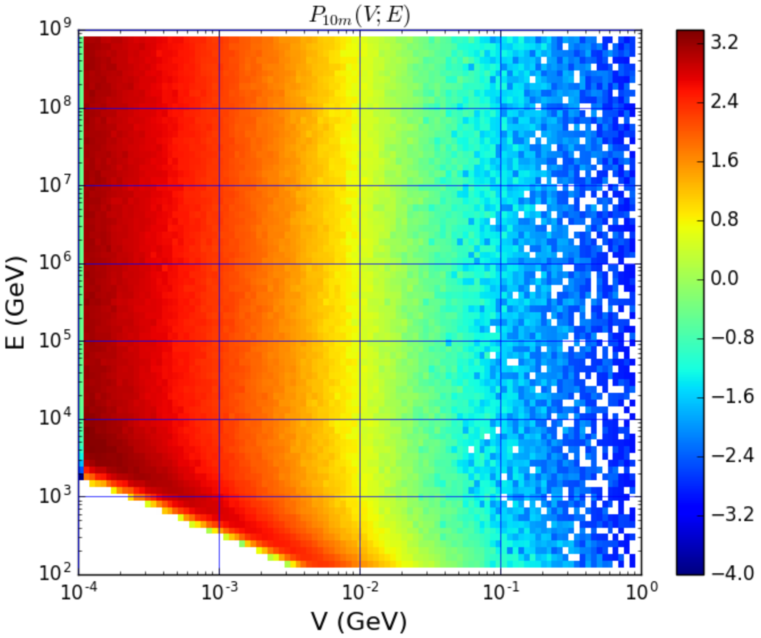


Figure 6.12: Probability distribution function built from PROPOSAL simulated losses of segment length $X = 10$ m binned along energy loss dE. The minimum energy loss produced by PROPOSAL causes the diagonal cut off line, except at higher muon energies $> 10^3$. If a larger range of V was chosen the diagonal line would continue and affect higher muon energies.

The consequence of the zero bins in the energy loss pattern is that if the PDF does not properly account for the zero bin and merely refers to the smallest possible energy loss (dE_{min}) it does not return the accurate probability for that zero loss, and results in poor resolution or inconsistent energy reconstruction distributions. Knowing this now, a PDF could be constructed that has a range that does properly account for zero energy losses, in this work we describe ways to remove the zero energy losses.

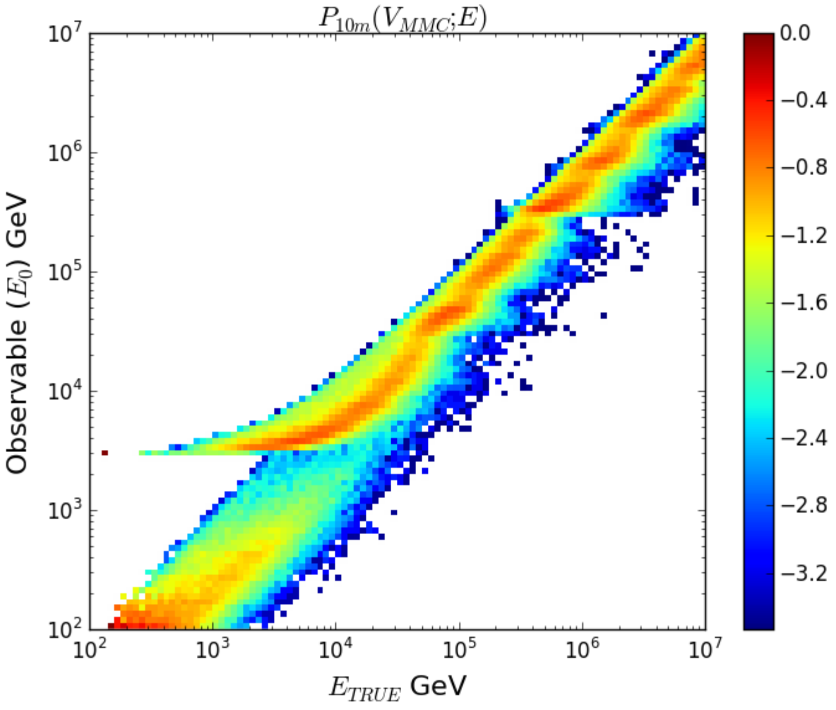


Figure 6.13: Edepillim reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE}, using the muon energy loss pattern of true losses (MMC) with the fractional energy loss implementation of Edepillim. The feature at $E_0 = 3 \times 10^3$ GeV is a consequence of the fractional PDF in Figure 6.12 having low probabilities for zero energy losses (which correspond the the minimum possible energy loss dE_{min}).

6.3.1 Continuous Energy

In order to produce consistency between the PDF and the simulation event's energy losses and to account properly for zero stochastic losses in the PDF, the continuous energy losses need to be taken into account. There is a consistent energy loss of 200 MeV/m for muons due to ionisation. This loss can be accounted for in the production of the PDF by taking the difference between the initial energy of the muon and the final energy after distance X, $dE = E_{initial} - E_{final}$, in PROPOSAL. The PDF as shown in Figure 6.14 for segment length $X = 10\,\text{m}$ has a more uniform shape with a minimum energy loss of $dE = 2\,\text{GeV}$, except at higher energies, $E > 10^5\,\text{GeV}$, where due to the small segment length there are still cases of no stochastic loss occurring and not all continuous loss is accounted for, due to limitations in PROPOSAL. This is an effect of PROPOSAL, but it will not impact the final results of the PDF as it will be smoothed out with the interpolated spline fitting (see Appendix C).

Millipede does not reconstruct the continuous losses in the energy loss pattern, but as we expect it is a fixed number for the segment length of the muon travelling in the medium, the correct energy can be added to every energy loss bin in the Millipede reconstructed energy loss pattern. Also it can be added to the true loss patterns from simulation to have consistency between the true energy losses and the PDF predicting these losses. Accounting for continuous energy creates a mapping effect of every zero energy loss bin corresponding to a minimum filled threshold bin in the PDF, resulting in consistency between the energy loss pattern and the PDF. All zero energy losses in the reconstructed and true energy loss pattern will now correspond to the threshold energy $dE_{threshold} = 0.2 \times X\,\text{GeV}$ in the PDF, where X is the segment length, and these losses will now be referred to as threshold energy losses.

The effect of accounting for the continuous energy, and thus accounting for zero stochastic losses in the PDF, is shown in the reconstruction distribution in Figure 6.15. There is a large improvement in the Edepillim reconstruction using the true energy losses in Figure 6.15 compared to the previous reconstruction in Figure 6.2; the correlation between the observed reconstructed energy and the true muon energy is clear, with a smaller spread in the reconstructed energies. The improvement in resolution is seen to be

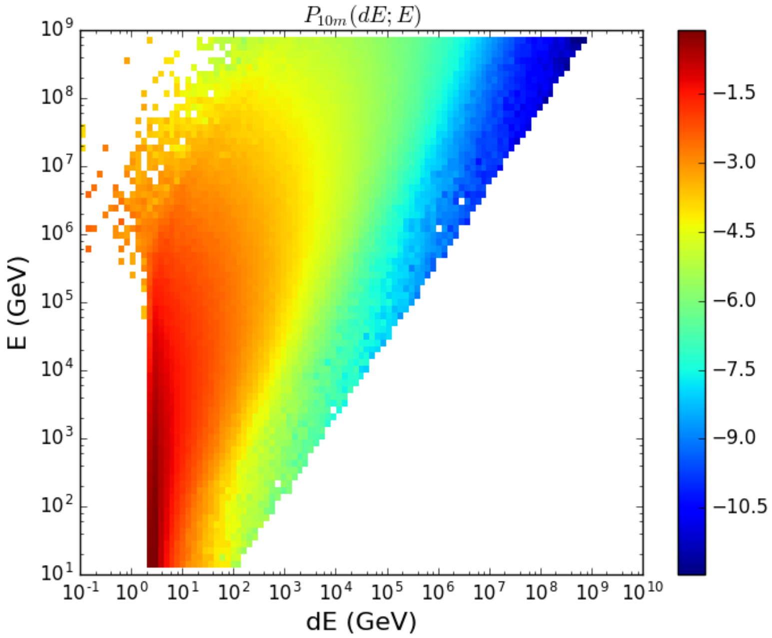


Figure 6.14: Probability distribution function built from PROPOSAL simulated losses using $dE = E_{initial} - E_{final}$ to account for continuous losses of segment length $X = 10\,\mathrm{m}$ binned along energy loss dE. The minimum energy loss produced by PROPOSAL causes the vertical cut off line, except at higher muon energies, above 10^5, where the small segment length means there are cases where there are no stochastic losses, just continuous losses that are not all accounted for, due to limitations in PROPOSAL. Theses effects are smoothed out in the interpolated fitter.

consistent across all energies as seen in the reconstruction resolution in Figure 6.16 where the energy reconstruction on true losses ($P_{10m}(a)(dE_{MMC};E)$) has improved compared to the initial results ($P_{10m}(dE_{MMC};E)$) with a resolution that is better than TE. This reconstruction resolution is now more similar to that of Chapter 5, which we expect since both energy loss patterns now contain true energy losses.

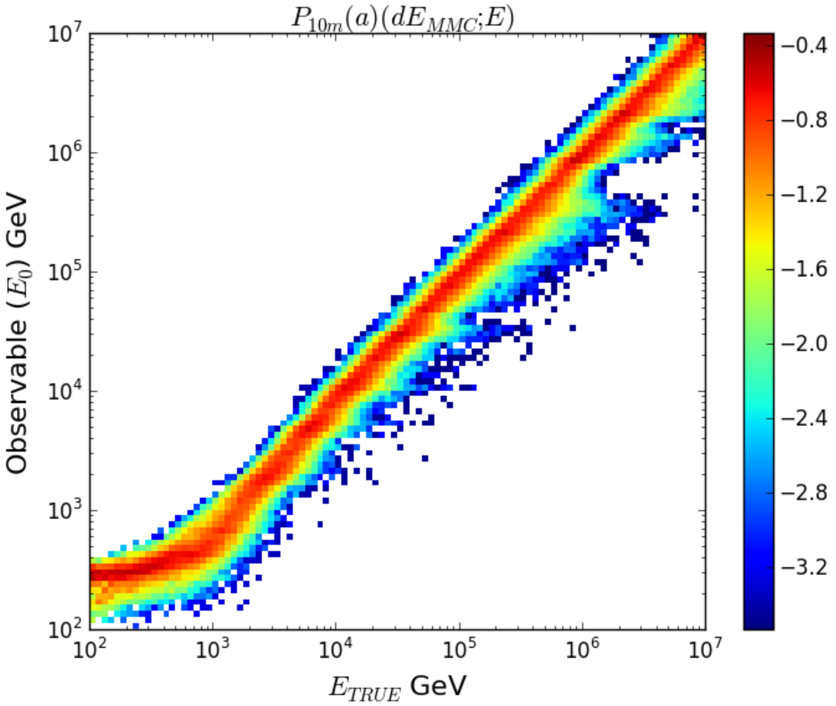


Figure 6.15: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for Edepillim reconstruction using the true losses pattern (MMC) with a continuous energy loss of 0.2 GeV added to all bins.

The energy reconstruction on Millipede losses ($P_{10m}(a)(dE_{Mill};E)$)) in Figure 6.16 has no noticeable change from its resolution in Figure 6.1. This is because other effects in the Millipede energy loss pattern overshadow any possible improvements from accounting for continuous energy. The next section will explore these other effects, but continuous energy losses will still be accounted for in the rest of this chapter as it ensures consistency between the PDF and the energy loss pattern used in the reconstruction.

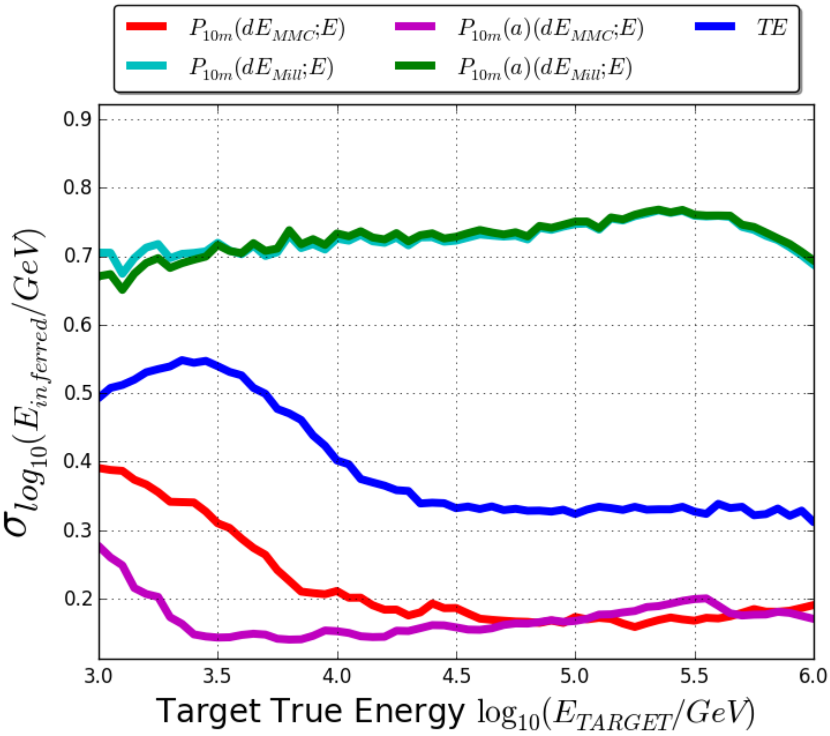


Figure 6.16: The ALF reconstruction resolution at target true muon energies, for the Edepillim reconstruction method on the true losses (MMC) and the reconstructed losses (Mill), for the initial testing of Edepillim, $P_{10m}(dE;E)$, and Edepillim with continuous energy losses added to the energy loss pattern, $P_{10m}(a)(dE;E)$. The inclusion of the continuous losses improves the resolution using the true energy loss pattern, but has no real impact when using reconstructed losses. The standard method for IceCube, the truncated mean energy (TE) reconstruction, is shown for comparison.

6.4 Improving Energy Losses

In order to improve the resolution of the Edepillim energy reconstruction method we need to improve the accuracy of the Millipede reconstructed energy loss patterns. In this section we explore removing possibly inaccurate large energy losses from outside the detector boundary.

6.4.1 Limiting Energy Losses to the Detector Volume

As can be seen in Figures 6.6 and 6.7 there are cases where Millipede reconstructs a large loss outside of the detector volume. This is caused by minuscule signals from edge DOMs that are most likely noise or small energy losses. Due to its maximum likelihood method and few physical restrictions on the likelihood of such a loss, Millipede tends to favour a distant large energy loss to explain these signals. This energy loss can often be many orders of magnitude greater than the true muon energy of the event. The overall effect of this can be seen in Figure 6.17, with a large tail present in the sum of losses towards higher values of $\sum dE_{Mill}$ than should be seen for the given $\sum dE_{MMC}$.

In order to minimise the effect in the energy reconstruction these inaccurate losses need to be removed. In order to allow for only losses that are more likely to be accurate, a cut was made on energy losses to include only those that occur within the detector volume. The detector volume was defined to be a cylinder with a height of 1050 m and radius of 575 m. By ensuring only such defined contained losses are used, all events of the simulation will have a varying path length (L) depending on their direction of travel in the detector. This will be further addressed in Section 6.6.2.

The improvement that containing the losses to the detector volume has can be seen in Figure 6.18 where only losses in the detector volume are included in the sum for both true and reconstructed energy losses. This results in a better correlation between the total of the true and reconstructed losses than was seen in Figure 6.17 where all losses were included, with less of a tail tending towards higher values of $\sum dE_{Mill}$.

The reconstruction distribution when using only losses in the contained detector volume has a great improvement for the Edepillim reconstruction using Millipede losses as shown in Figure 6.19, mostly on the events that were previously seen in the top left sec-

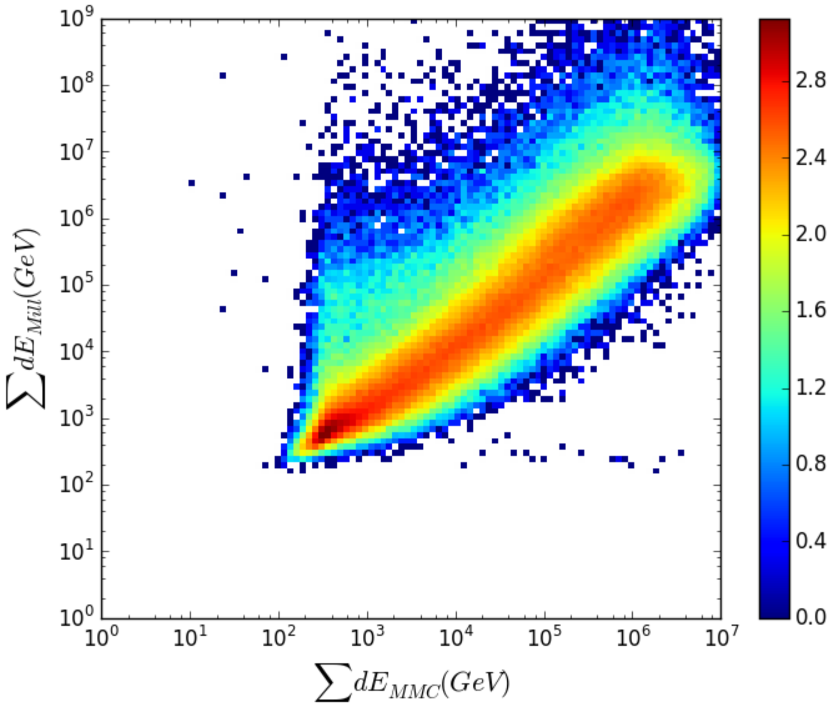


Figure 6.17: The sum of the losses in the true energy loss pattern $\sum dE_{MMC}$ vs the sum of the losses in the Millipede reconstructed energy loss pattern $\sum dE_{Mill}$ on a log scale histogram. While the bulk of events are consistent, there are outlier events that do not have a reconstructed total energy loss that is consistent with the total true energy loss, with some events having a $\sum dE_{Mill}$ that is several orders of magnitude higher than its $\sum dE_{MMC}$.

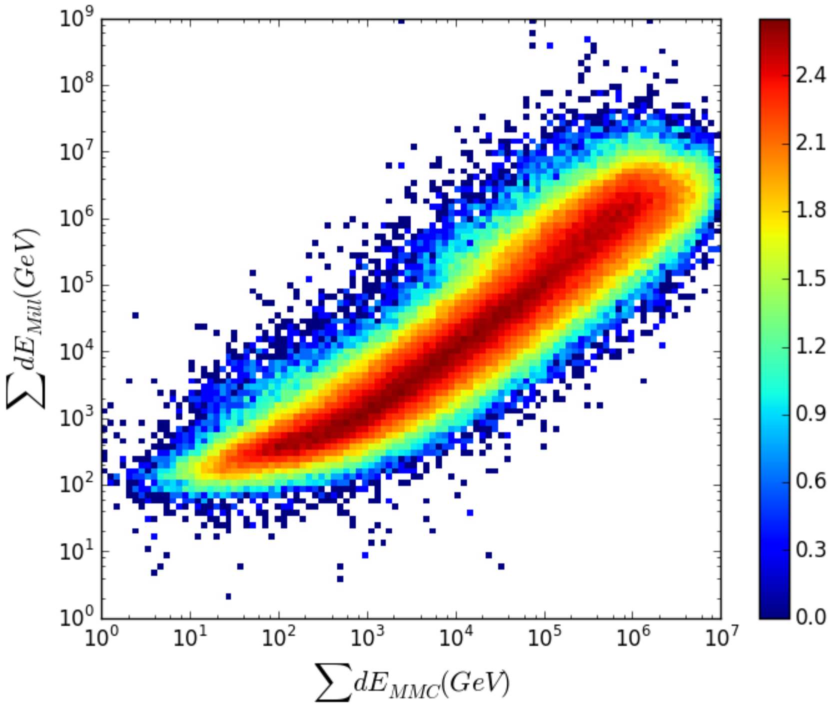


Figure 6.18: The sum of the losses contained within the detector volume in the true energy loss pattern $\sum dE_{MMC}$ vs the sum of the losses in the Millipede reconstructed energy loss pattern $\sum dE_{Mill}$ on a log scale histogram. There are very few outlier events for the reconstructed total energy loss compared to Figure 6.17, with the contained losses also reducing the lower limit of the sum of losses for both reconstructed and true losses, with the most noticeable being a shift to the left for the total true losses.

tion of Figure 6.3 that were reconstructed at much higher energies than their true muon energy. As the energy loss pattern such as in Figure 6.6 would no longer include the large energy loss, the physical constraint ($E_0 > \sum dE$) on the reconstructed energy has been lowered, and the Edepillim reconstruction can now operate with greater resolution. This is seen in the reconstruction resolution in Figure 6.21 with an overall improvement in the energy reconstruction with Millipede losses at all energies ($P_{10m}(b)(d\vec{E}_{Mill}; E)$) compared to the previous reconstruction resolution ($P_{10m}(a)(d\vec{E}_{Mill}; E)$).

There are also very small changes in Figure 6.20 of the reconstruction distribution of Edepillim using the true energy losses contained in the detector volume compared to the previous reconstruction resolution in Figure 6.15. These small fluctuations are caused by there now being fewer losses in the energy loss pattern in the reconstruction for $P_{10m}(b)(d\vec{E}_{MMC}; E)$ compared to the previous reconstruction $P_{10m}(a)(d\vec{E}_{MMC}; E)$, which as explained in Chapter 5, the more information in the energy loss pattern the better the resolution. Therefore as $P_{10m}(b)(d\vec{E}_{MMC}; E)$ has is slightly less information the reconstruction resolution fluctuates.

6.5 Threshold Energy Losses Effect on Edepillim Reconstruction

Within the Millipede reconstructed energy loss pattern there is an abundance of threshold losses: energy losses with zero energy loss or extremely small losses. When accounting for continuous energy losses (Section 6.3.1) these losses are now at the minimum of continuous energy $0.2 \times X$ GeV thus correspond to losses at the edge of the PDF. The effect that a majority of threshold losses have on the muon reconstructed energy is demonstrated in the following series of likelihood spaces, where we artificially change some true energy losses into threshold losses. In Figure 6.22 the likelihood space of the energy reconstruction of the true losses, with continuous losses accounted for, and detector boundary conditions applied, is shown with a smooth minimum found near the true muon energy. In Figure 6.23 for the same muon event only 50% of the highest true losses are kept in the energy loss pattern. The remaining losses are set to be threshold losses, $dE_{threshold} = 0.2 \times X$ GeV, where X is the segment length. Edepillim has then reconstructed a different E_0 than

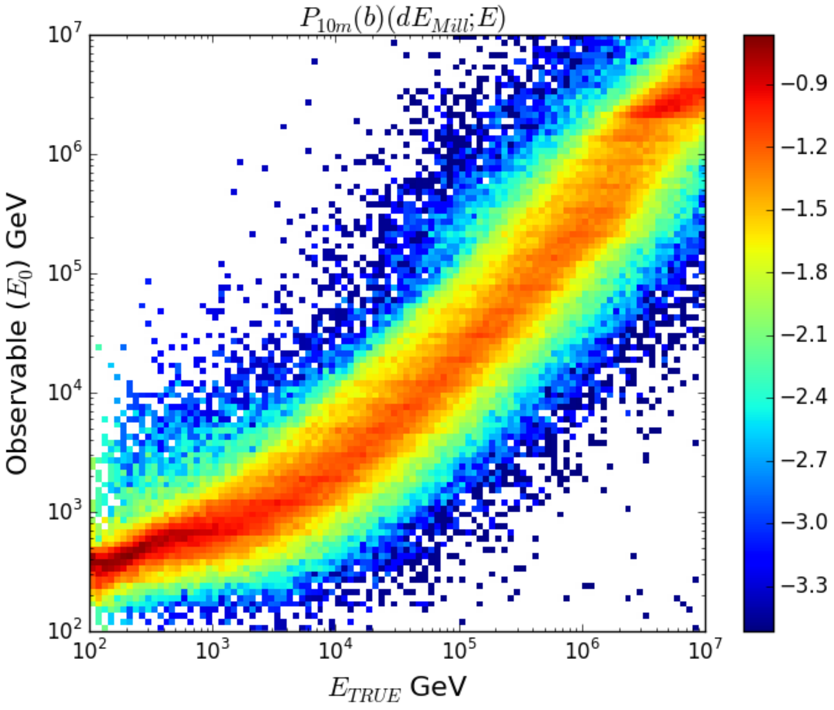


Figure 6.19: Edepillim reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy losses pattern (Mill) with continuous energy losses added to the energy loss pattern and only including losses contained within the detector volume. The reconstruction has improved with fewer events reconstructed to higher energies (top left corner) compared to the previous results in Figure 6.3.

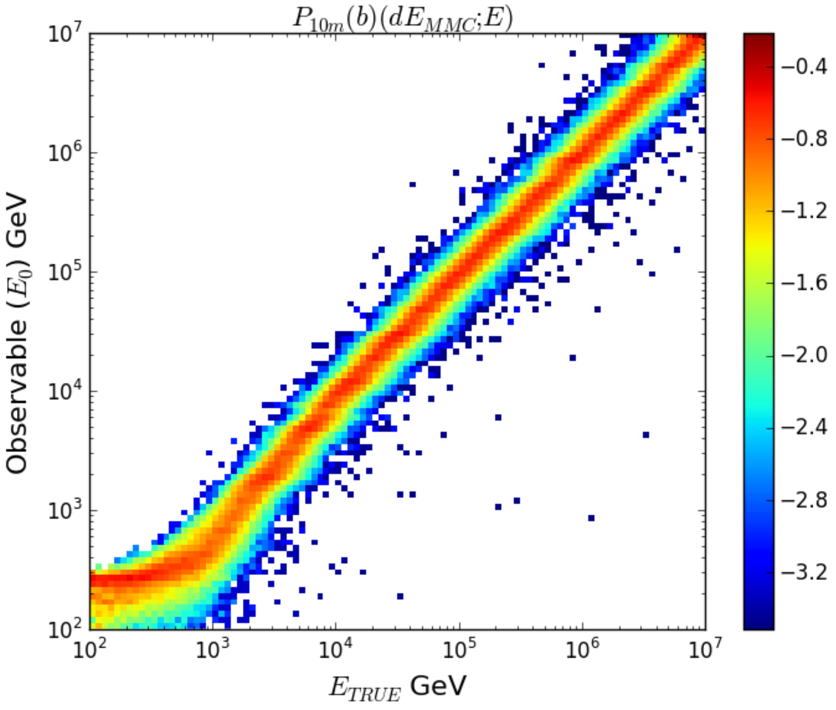


Figure 6.20: Edepillim reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for muon energy losses pattern of true losses (MMC) with continuous energy losses added to the true energy loss pattern and only including losses contained within the detector volume. There are now fewer energy losses in the pattern than in the previous results in Figure 6.15 which results in minor fluctuations to the reconstruction resolution.

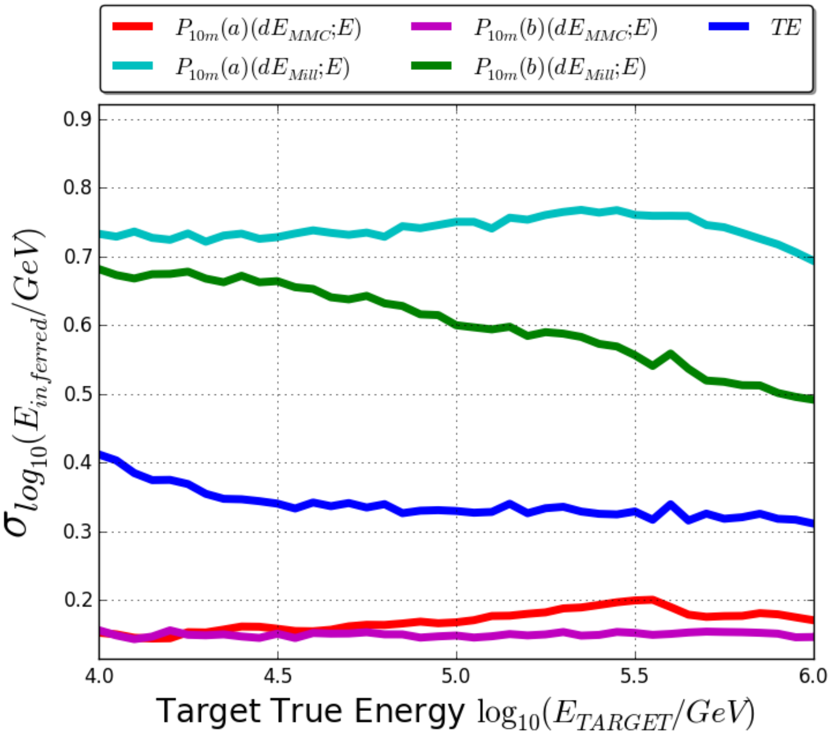


Figure 6.21: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the true losses (MMC) and the reconstructed losses (Mill) for the previous reconstructions with continuous energy losses added to the energy loss pattern, $P_{10m}(a)(d\vec{E};E)$, compared to using continuous losses and only including losses contained within the detector volume, $P_{10m}(b)(d\vec{E};E)$. With the cut on detector volume the resolution of Edepillim reconstruction on reconstructed energy losses improves. The IceCube standard reconstruction method, truncated mean energy (TE), is shown for comparison.

in Figure 6.22, with the new E_0 moving away from the true muon energy (E_{TRUE}) and approaching the sum of the energy losses in the pattern ($\sum(d\vec{E})$). This trend continues if more energy losses of the event are replaced with threshold losses, until Figure 6.24 where only 30% of the true losses remain and the Edepillim reconstructed energy is now simply the sum of all the energy losses. This type of likelihood space is defined as a 'stuck' likelihood, as its reconstructed energy is stuck at the first possible energy given the physical limitation that the muon energy must be greater than the sum of the energy losses. An event with a 'stuck' likelihood does not have a working Edepillim energy reconstruction as there is no minimisation taking place. This 'stuck' answer could be the correct energy in the case of a muon running out of energy along its path length, in which case Edepillim is working, but there is no information supplied to the likelihood about the position of the threshold losses along the path length to help with this conclusion, i.e. if all the threshold losses were at the end of the energy loss pattern, this would be a good reconstruction, but the likelihood does not know where along the muon track the loss occurred.

This trend of the reconstruction result approaching the sum of energy losses when a decreased percentage of losses is kept happens for all muon events. The level of percentage of energy losses remaining when the reconstructed energy is the sum of all energy losses is different for different individual muons events as the remaining energy losses will still have influence on the likelihood space, depending on their probabilities. But the overall reconstruction resolution is shown in Figure 6.25 with the trend becoming noticeable at 50% and dominating at 40% to 10%. Figure 6.25 also includes the reconstruction resolution if one takes $E_0 = \sum d\vec{E}_{MMC}$ which is the value the Edepillim method returns when the likelihood is 'stuck'. The resolution for the sum of the losses is between the 40% to 30% results. Note that the 10% limit is much worse than the sum of losses because only 10% of the losses remain and so the summation of losses $\sum d\vec{E}_{MMC}(10\%)$, will be less than the sum for all the losses in the pattern, $\sum d\vec{E}_{MMC}$.

This behaviour is also the same if the total sum of all losses is conserved, that is if the lower losses are set to threshold losses, but their remaining energy is redistributed to the higher losses. This mimics the tendency of Millipede to reconstruct large losses in a few energy loss bins with many threshold energy bins. The threshold losses still overwhelm

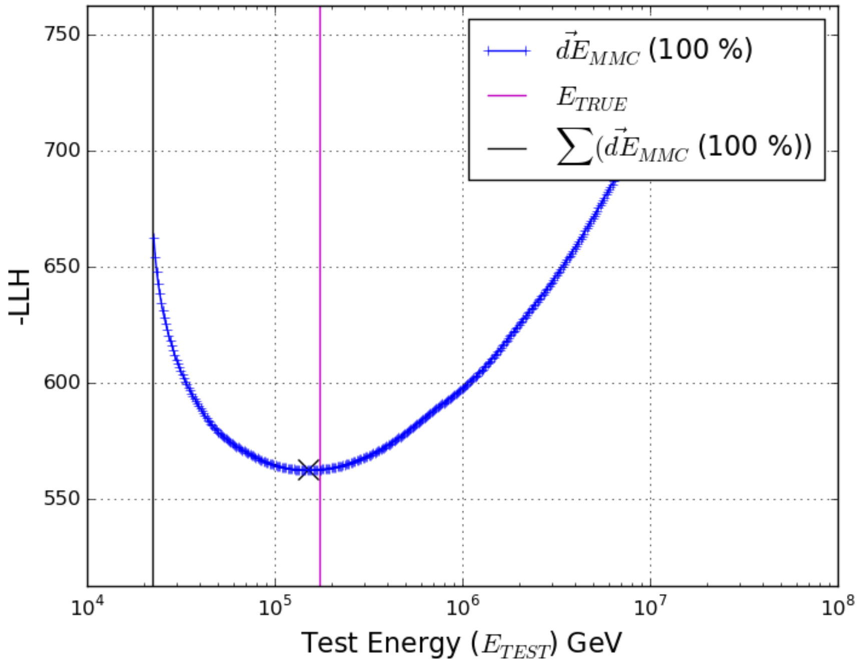


Figure 6.22: The Edepillim likelihood space for a muon event using the true energy losses $d\vec{E}_{MMC}$. The minimum of the likelihood space, i.e. the reconstruction observable is shown by the cross at $E_0 = 1.5 \times 10^5$ GeV, as well as the true muon energy (*pink*) and the sum of all the energy losses in the pattern (*black*).

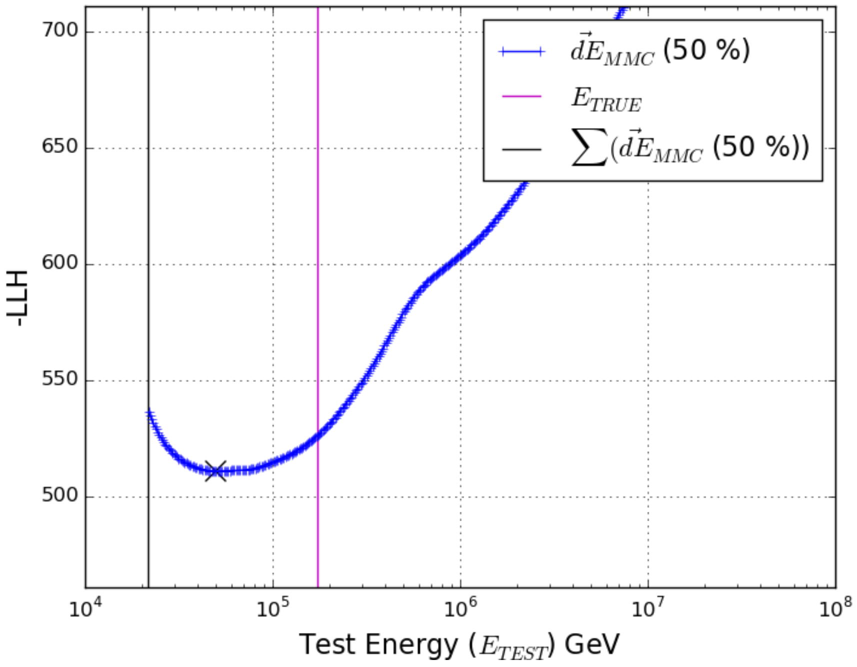


Figure 6.23: The Edepillim likelihood space for a muon event using the highest 50% of the true energy losses $\vec{dE}_{MMC}$, with the remaining 50% of losses being set to threshold losses. The minimum of the likelihood space, i.e. the reconstruction observable is shown by the cross at $E_0 = 5 \times 10^4$ GeV, as well as the true muon energy (*pink*) and the sum of all the energy losses in the pattern (*black*).

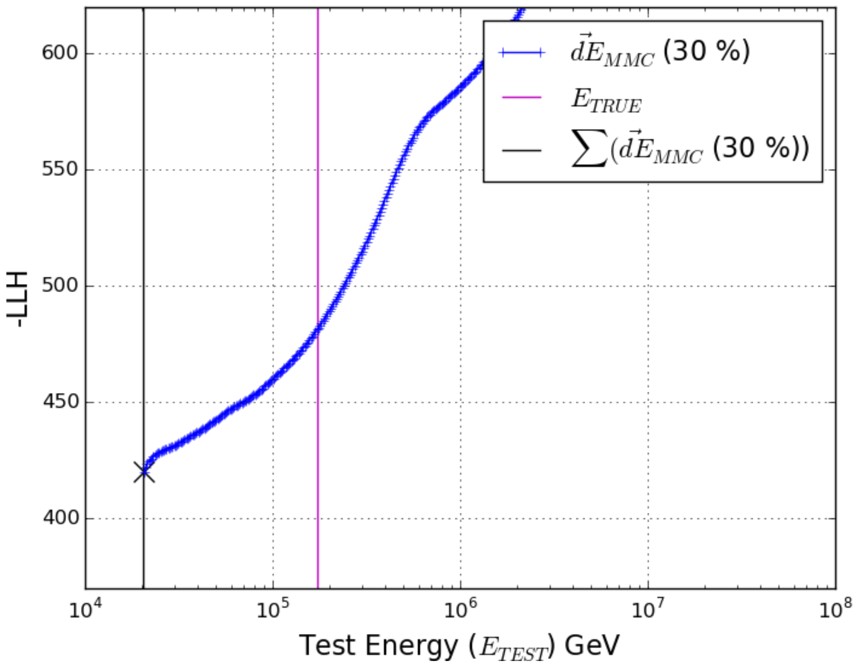


Figure 6.24: The Edepillim likelihood space for a muon event using the highest 30% of the true energy losses $d\vec{E}_{MMC}$, with the remaining 70% of losses being set to threshold losses. The minimum of the likelihood space, i.e. the reconstruction observable is shown by the cross at $E_0 = 2 \times 10^4\,\text{GeV}$, as well as the true muon energy (*pink*) and the sum of all the energy losses in the pattern (*black*). In this case the likelihood is 'stuck' at the first possible muon energy allowed due to the physical limitation of it being greater than the sum of the losses.

the likelihood leading to the reconstructed energy being the sum of the losses, but since the sum of losses is now conserved, the resolution is never poorer than the simple sum of the losses method, as shown in Figure 6.26.

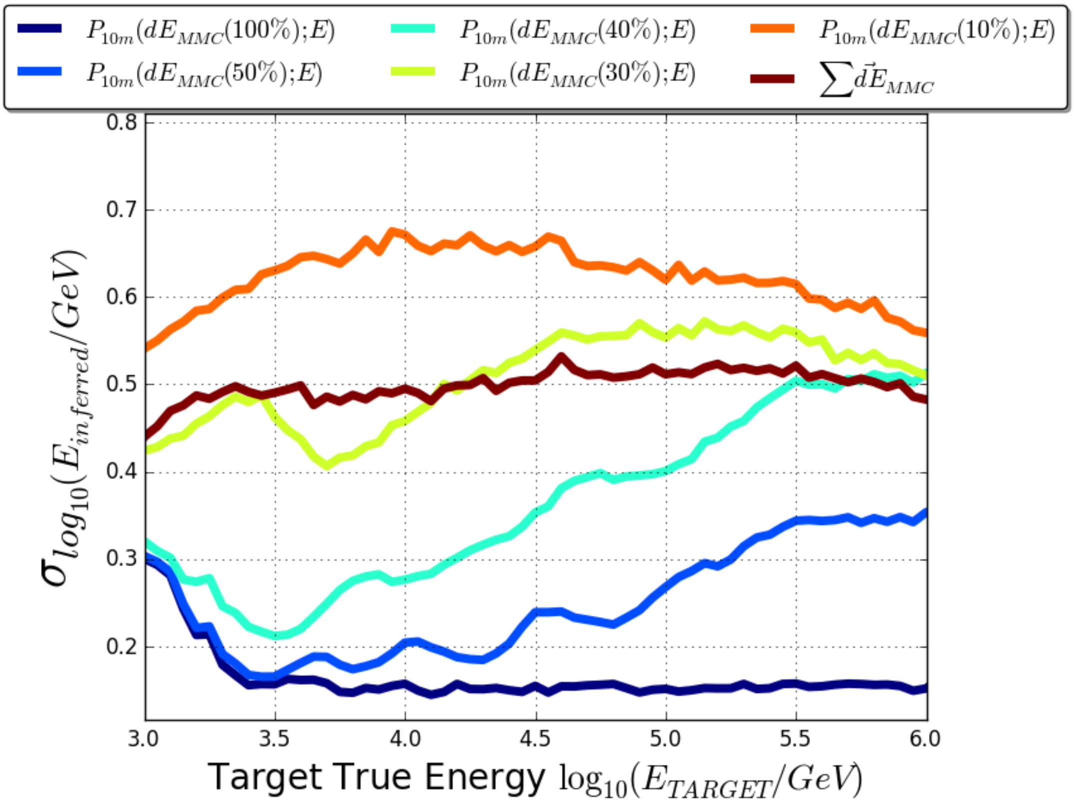


Figure 6.25: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the true losses (MMC) with a percentage of the highest energy losses kept and the remaining energy losses set to the threshold loss $dE_{threshold} = 2\,\mathrm{GeV}$. Shown for comparison is the reconstruction resolution if the reconstructed energy is given by the sum of the energy losses, $\sum d\vec{E}_{MMC}$.

The threshold energy loss trend is detrimental to the Edepillim reconstruction method as all of the muon events within this simulation have an abundance of threshold energy losses in their Millipede reconstructed energy loss pattern. In Figure 6.27 the events are shown with the percentage of their above-threshold energy losses of their energy loss pattern. There are a large number of events that sit below 40% i.e. only $0-40\%$ of their losses are above the threshold value. In total 74% of the events have energy loss patterns where their above threshold losses are between $0-40\%$. This results in a large number

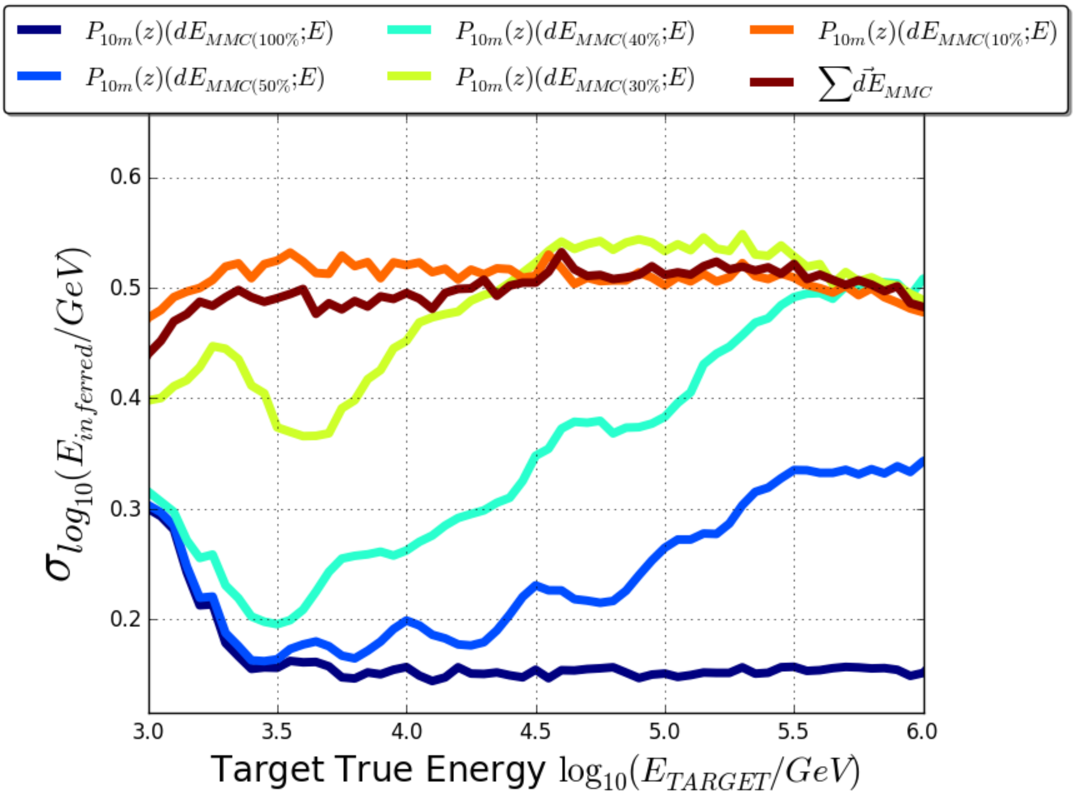


Figure 6.26: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the true losses (MMC) with a percentage of the highest energy losses kept and the remaining energy losses set to the threshold loss $dE_{threshold} = 2\,\text{GeV}$, with the rest of their energy loss redistributed to the highest losses that are kept. Shown for comparison is the reconstruction resolution if the reconstructed energy is given by the sum of the energy losses, $\sum d\vec{E}_{MMC}$.

of cases where the Edepillim method is not working as no clear minimum is found, just the sum of the energy losses as demonstrated in Figure 6.25. This can be easily seen in Figure 6.28 where the reconstruction resolution of Edepillim using Millipede losses is almost identical to that from using the sum of the energy losses as the reconstructed energy.

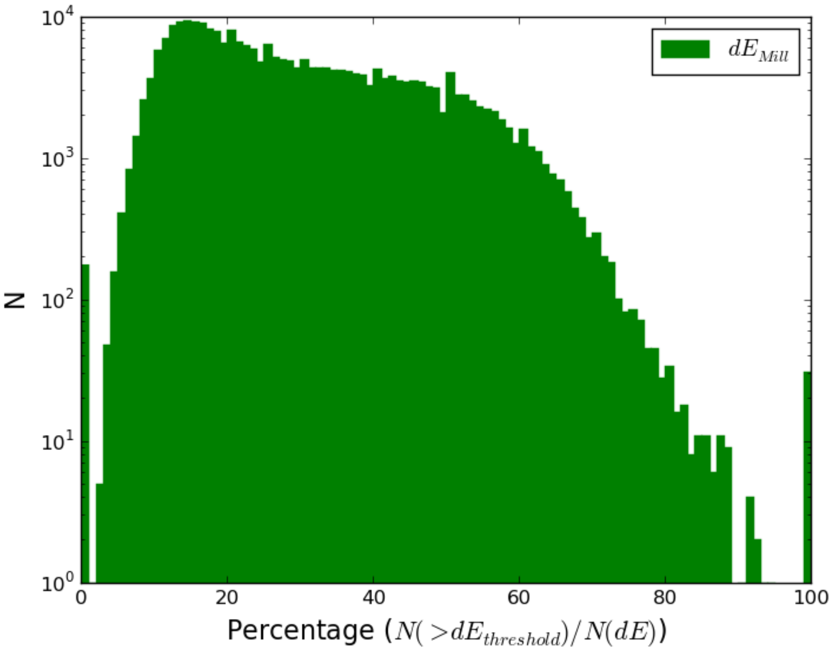


Figure 6.27: Histogram of the percentage of energy losses above the threshold energy in each muon event's reconstructed energy loss pattern. The events below approximately 40% will have too many threshold losses and will not be performing a correct Edepillim reconstruction.

For the 74% of events where the percentage of above-threshold energy losses is below 40%, the abundance of threshold losses in these events lead them to reconstruct the energy as the sum of the energy losses, i.e. a 'stuck' likelihood space. Because of the large number of these 'stuck' events the reconstruction is working only at the same resolution as taking the sum of all losses seen, $\sum dE$, as the reconstructed observable. This is not a working reconstruction as none of the power of the Edepillim method is being utilised, instead the threshold losses are causing the reconstruction to become 'stuck' at the sum of the energy losses, which is the first possible energy that is allowed because of the physical limitations enforced in Edepillim.

The abundance of threshold losses reconstructed by Millipede is a flaw in the reconstruction that as discussed in this section causes significant problems in the Edepillim energy reconstruction. This flaw does not affect the total of the energy losses reconstructed by Millipede as Figure 6.18 shows that the sum of the reconstructed energy loss correlates with the sum of the true losses.

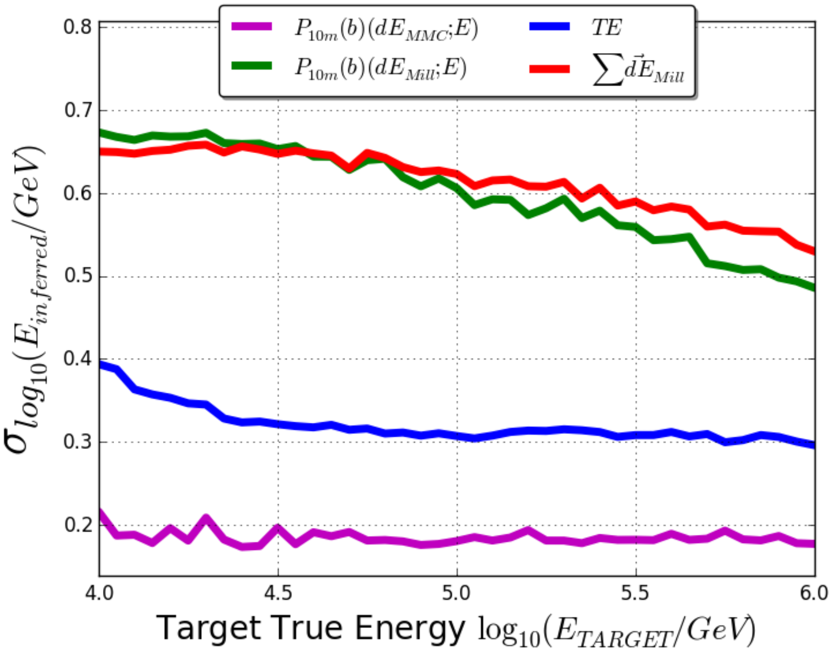


Figure 6.28: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the true losses (MMC) and Millipede reconstructed losses (Mill). Shown for comparison is the reconstruction resolution for TE (TE) and for the case that the reconstructed energy is given by sum of the energy losses $\sum d\vec{E}_{MMC}$. The resolution for the sum of the energy losses is at the same level as the Edepillim reconstruction using Millipede losses indicating that the Edepillim reconstruction is often reconstructing the muon energy to the sum of the energy losses.

6.5.1 Segment Length Rebinning

In order to achieve a better Edepillim reconstruction resolution the accuracy of the energy loss pattern needs to be improved. For the best resolution the improvement needs to be in the estimate of large energy losses, that are often in the true case a cluster of energy losses, and to reduce the number of threshold losses in the pattern so they do not dominate the likelihood space, leading to stuck likelihoods.

The best way for improving the losses reconstructed by Millipede, which will be shown in this section, is to continue to reconstruct with Millipede at $X = 10\,\mathrm{m}$, but afterwards to create new energy loss patterns with increased segment length by addition of adjacent $X = 10\,\mathrm{m}$ energy loss bins. By combining the bins after the Millipede reconstruction has been done, the new segment lengths result in fewer energy threshold bins and a closer approximation to the true losses, as will be shown in this section. An example of an event's addition of adjacent energy loss bins to rebin the energy losses to artificially increase the segment length can be seen in Figure 6.29, which was originally the energy loss pattern shown in Figure 6.6. As the segment length is rebinned to increasing sizes the energy losses become closer to the true losses after the same rebinning.

Of course every event has a different path length due to its trajectory through the detector. When there is a shorter path length, rebinning results in a reduced number of losses such as in Figure 6.30. There is also a missing loss at the end of the pattern since, if an event was not long enough such that the final losses would then not fit into a full segment length bin, those remaining energy losses were removed from the energy loss pattern.

The overall effect of using larger combined segment lengths is shown by comparing Figures 6.31 and 6.32. In Figure 6.31 the comparison of each bin's true and Millipede energy loss for all events in the simulation is shown with an overall large spread of the energy losses. The large variation of the reconstructed loss with respect to the true loss means that the majority of energy losses are not correct. This will affect the Edepillim reconstruction as it relies on knowing the energy losses as accurately as possible. There are also a large number of threshold energy losses, represented by the horizontal line at $dE_{Mill} = 2\,\mathrm{GeV}$.

After the segment length rebinning, combining ten segment lengths of $X = 10\,\mathrm{m}$ to make a segment length of $X = 100\,\mathrm{m}$, Figure 6.32 shows that there is a more clear correlation between the true energy losses and the reconstructed losses, with a reduced width in the possible reconstructed loss, making it more reliable for use in an energy reconstruction. There are still some threshold losses, represented in Figure 6.32 as the horizontal line at $dE_{Mill} = 20\,\mathrm{GeV}$. This is more clearly seen in Figure 6.33, where, for a

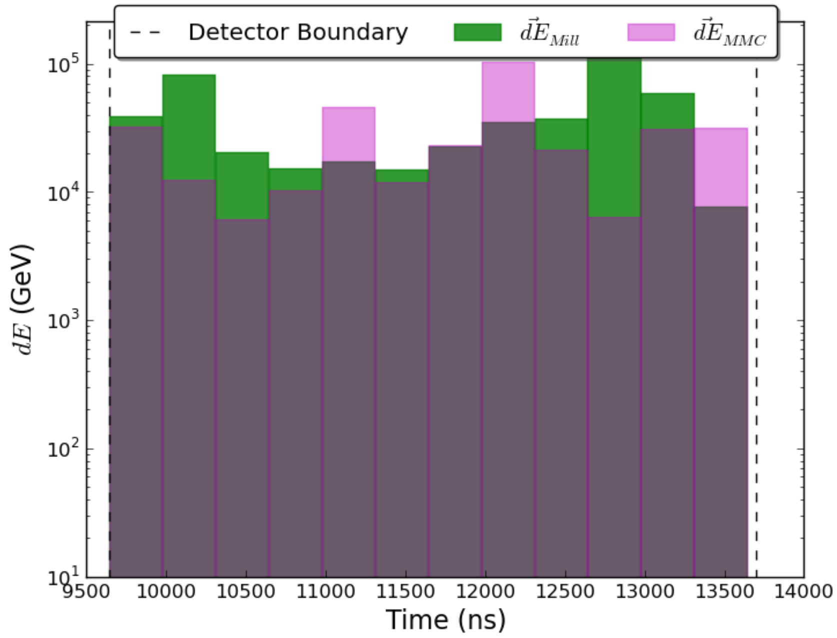


Figure 6.29: A simulated muon's energy loss pattern for its rebinned segment length $X = 100\,\text{m}$ for true losses $d\vec{E}_{MMC}$ (*pink*) and its reconstructed losses $d\vec{E}_{Mill}$ (*green*). The dashed lines show the detector boundary for the track. This event's energy loss pattern was originally that in Figure 6.6 but here with the energy losses being contained there are no large energy losses outside the detector volume and with rebinned segment lengths there is a pattern of more constant losses that resemble the true losses. It is also important to note that the pattern now contains no threshold losses that can dominate the reconstruction likelihood.

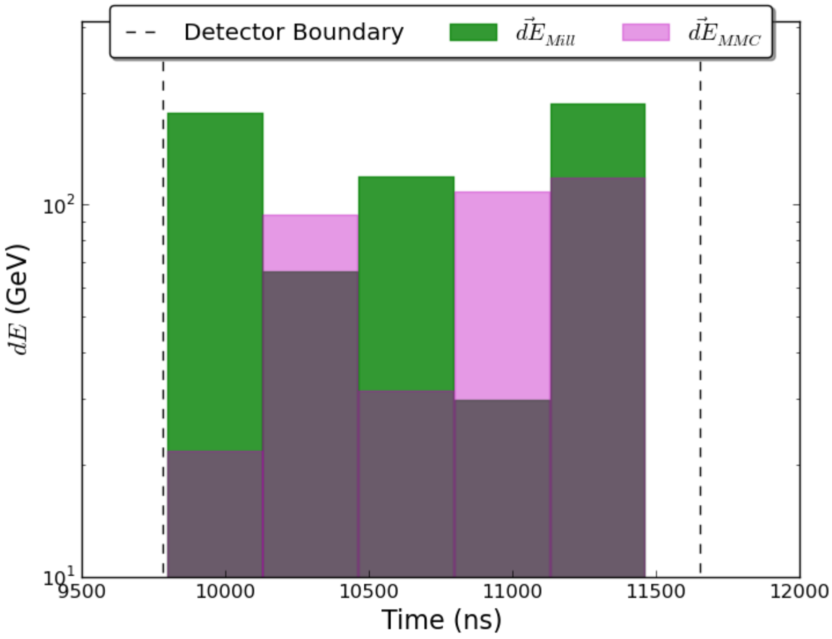


Figure 6.30: A simulated muon's energy loss pattern for its rebinned segment length $X = 100\,\mathrm{m}$ for true losses $\vec{dE}_{MMC}$ and its reconstructed losses $\vec{dE}_{Mill}$. The dashed lines show the detector boundary for the track. This event's energy loss pattern was originally that in Figure 6.5 but here all energy losses are within the detector boundary. There is a shortened energy loss pattern such that with the rebinned segment length there are only five energy losses.

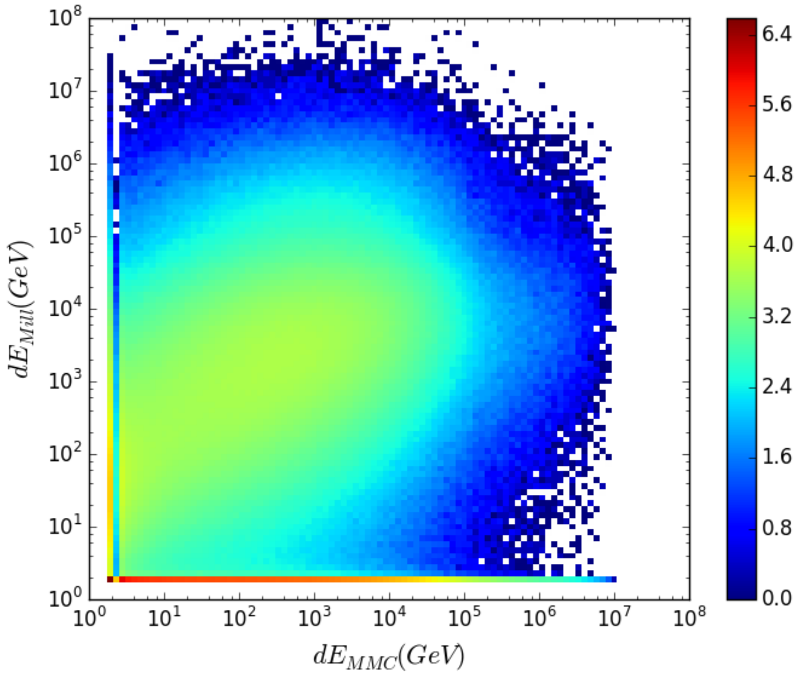


Figure 6.31: Comparison of the Millipede reconstructed energy loss to the true energy loss for segment length $X = 10\,\mathrm{m}$, for the log of the number of events. There is a large spread in the bin to bin comparison of energy losses, with a large number of threshold losses at $dE_{Mill} = 2\,\mathrm{GeV}$

specific true energy loss, the spread in the reconstructed energy loss can be seen, with a reduction in the spread for the rebinned segment length $X = 100\,\mathrm{m}$.

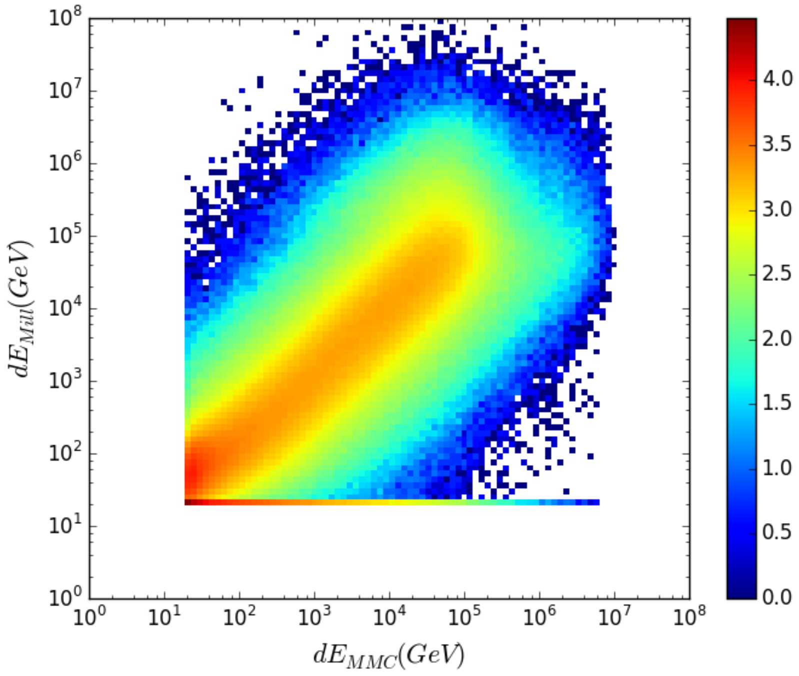


Figure 6.32: Comparison of the Millipede reconstructed energy loss to the true energy loss for rebinned segment length $X = 100\,\mathrm{m}$. The spread in the bin to bin comparison of energy losses is now not as large, and a clear connection between the true energy loss and the reconstructed Millipede loss can be seen. This is an improvement from Figure 6.31 where the correlation is not as clear.

The effect of larger segment length in this section is in contrast to that explored in Section 5.2.1, where in the true energy loss case increasing the segment length produces a worse energy resolution. The rebinning in the true case caused a loss of detail in the energy loss pattern; in the case of the Millipede losses considered here the rebinning averages the reconstructed losses closer to the truth. But there is a limit to the effectiveness of this technique, if the losses are rebinned in segments that are too large then the energy resolution will get worse. Thus there is expected to be an optimal rebinning after which there is a turn over in the resolution level. Figure 6.34 shows this turn over, increasing the segment length sees the resolution improve until $X > 200\,\mathrm{m}$ after which the resolution becomes progressively worse. At these track lengths the improvement in the Millipede

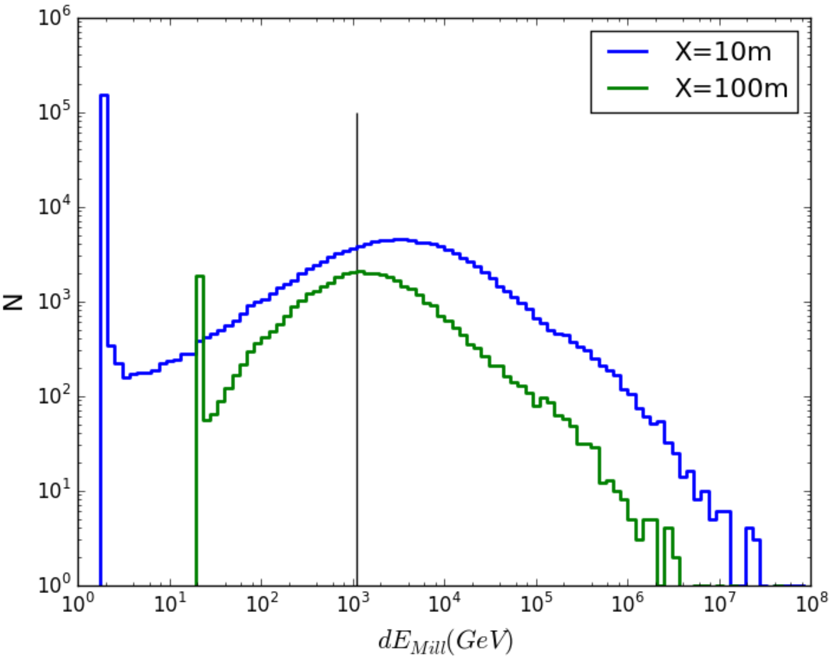


Figure 6.33: The distribution of the Millipede reconstructed energy loss for losses with true energy loss $dE_{MMC} = 10^3\,\text{GeV}$ for rebinned segment lengths of $X = 10\,\text{m}$ and $X = 100\,\text{m}$. The vertical black line shows the true energy dE_{MMC}. The spread in the bin to bin comparison of energy losses for $X = 100\,\text{m}$ is not as large as for $X = 10\,\text{m}$. Additionally there are fewer threshold events for $X = 100\,\text{m}$, shown by a reduction in the spike at threshold energies.

losses is obscured because too much detail has been lost in the subsequently obtained energy loss pattern for a good Edepillim reconstruction.

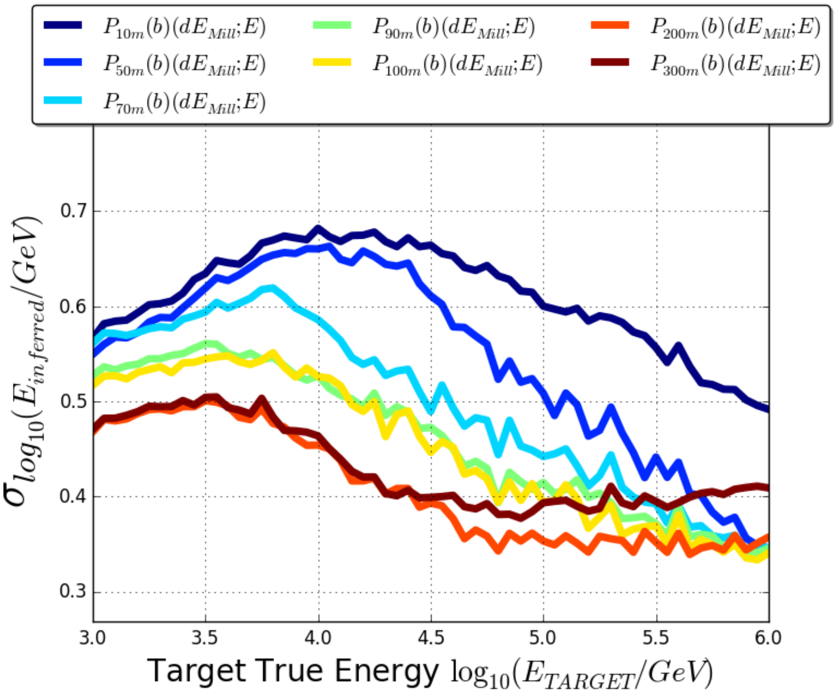


Figure 6.34: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with various rebinned segment lengths. The optimal segment length is $X = 200\,\mathrm{m}$ (reconstruction distribution shown in Figure 6.35), with the resolution worse at higher and lower rebinned segment lengths.

The reconstruction distribution for the optimal segment length is shown in Figure 6.35. The distribution has a reduced width compared with the reconstruction for segment length 10 m in Figure 6.19, and it also has a more consistent slope to the distribution which will improve the accuracy of energy reconstruction at a larger range of energies.

6.5.2 Threshold Energy Loss Cut

While the segment length rebinning reduces the number of threshold energy losses they are still present in the energy loss pattern. In Figure 6.36 there are only approximately 1.7% of events that have enough threshold losses for a stuck likelihood to be dominant, that is their percentage of above threshold losses is less than 40%. In Section 6.6.1 a selection criteria for removing any event with a stuck likelihood is explored. While the stuck

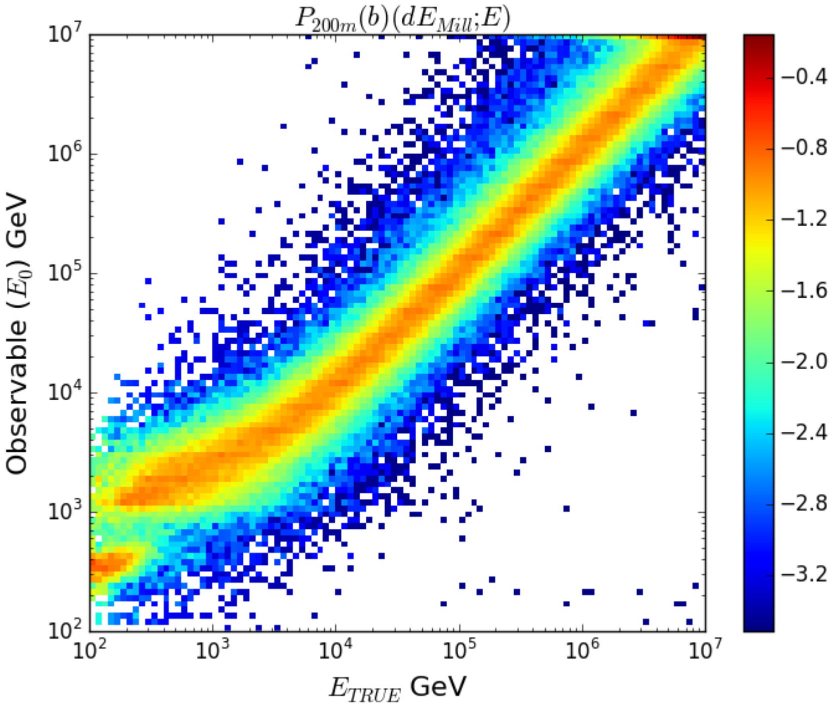


Figure 6.35: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy loss pattern (Mill) that has had segment lengths rebinned to $X = 200\,\mathrm{m}$. This is the optimal segment length for the best reconstruction resolution, for the muon simulations in IceCube considered here.

likelihood is no longer dominant at rebinned segment lengths there are still threshold losses present. It was found that for all the muon energy loss patterns for rebinned segment length $X = 100\,\text{m}$ approximately 12% of all the Millipede reconstructed losses in the simulation are threshold losses.

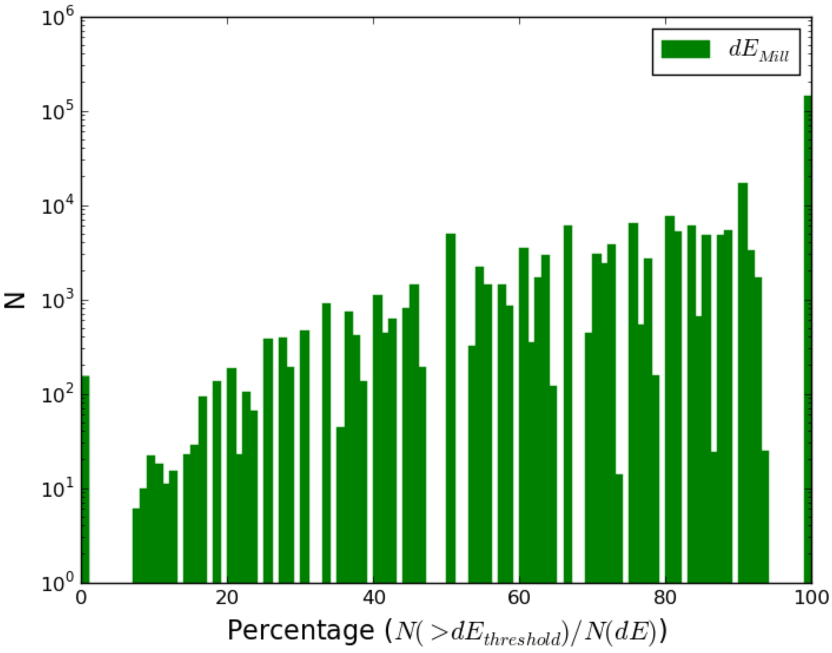


Figure 6.36: Histogram of the percentage of energy losses above the threshold energy in each muon event's energy loss pattern, for segment length $X = 100\,\text{m}$. The events below approximately 40% will have too many threshold losses and will not be performing a correct Edepillim reconstruction.

In order to reduce the impact of these threshold losses on the energy reconstruction a cut is introduced in the log likelihood calculation where any threshold loss was not allowed to contribute a probability to the likelihood. While this negates the workaround discussed in Section 6.3.1 of accounting for continuous losses, it is necessary to improve the resolution. But continuous losses should still be accounted for to maintain consistency as the PDF that is used is continuous, thus the energy losses in the pattern should still include a continuous contribution. With improved energy losses the threshold energy loss cut could be avoided and all energy losses included in the reconstruction.

The improvement in the resolution is shown in Figure 6.37 where various segment lengths are again tested. With the threshold cut, $X = 200\,\text{m}$ is no longer the optimal rebinning, instead there is overlap between $X = 50 - 200$ m with a slight loss in resolution at higher energies for $X = 200$ m. This cut puts the $X = 50 - 200$ m Edepillim reconstructions at the resolution level of $\sigma = 0.3 - 0.45$, whereas without the cut only segment lengths $X = 200 - 300$ m had resolutions in that range. Ideally you want to keep as many energy losses as possible in the pattern as shown in Section 5.2.1, to improve resolution. Therefore we will take $X = 100$ m as the optimal segment length rebinning.

It is also important to note that the reconstruction does not work if you ignore the threshold losses without first rebinning the segment length. This is because of the high percentage of threshold losses at lower segment lengths. Ignoring them does not provide enough information for the energy reconstruction and results in an inconsistent reconstruction distribution like the one seen in Figure 6.38.

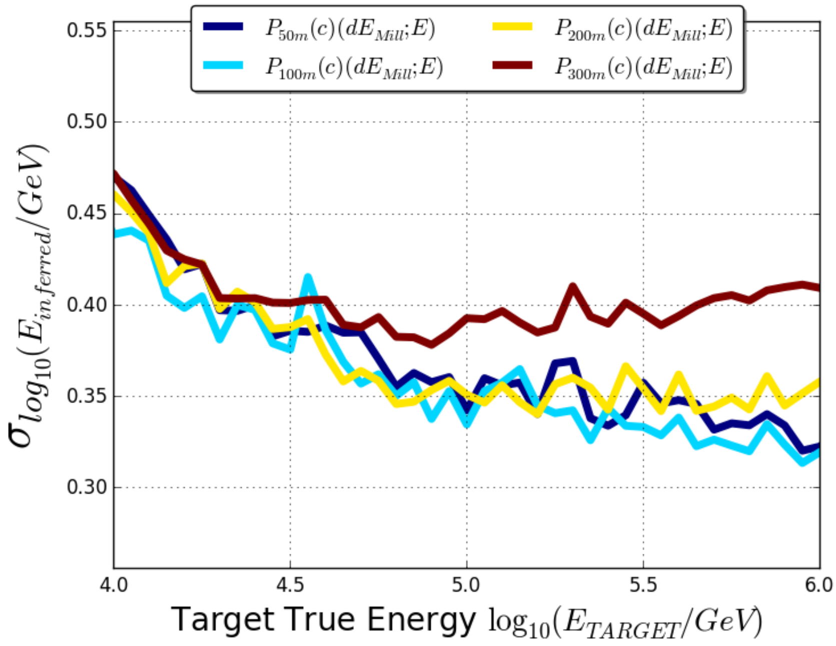


Figure 6.37: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with various rebinned segment lengths. All events have removed threshold energy losses form the likelihood calculation for the reconstruction. Above segment length $X = 50$ m the resolution is consistent until $X = 200$ m where the resolution begins to get worse.

In Figure 6.39 the reconstructed resolution of TE is compared to the the Edepillim reconstruction using a rebinned segment length of $X = 100$ m with threshold losses re-

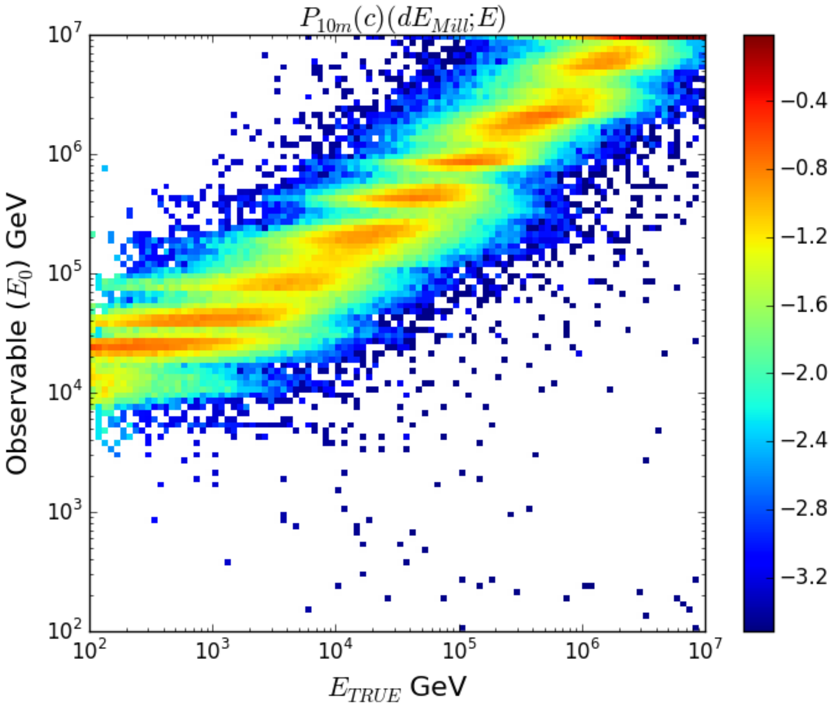


Figure 6.38: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy losses pattern (Mill) that use segment length $X = 10\,\text{m}$ and all threshold losses are ignored in the likelihood. This distribution does not have a good resolution value because of its large width, low slope and inconsistent peaks in slices along true energy.

moved from the likelihood, as well as compared to the results of the same reconstruction procedure on the true losses (MMC), as this is the expected optimal performance of Edepillim. With the improvement in the energy losses by rebinning the segment length and not allowing the threshold losses to contribute to the likelihood, the Edepillim energy reconstruction on Millipede events has begun to approach the TE resolution. The Edepillim reconstruction still has a poorer resolution than what can be achieved with the true losses but it is of course not going to reach that resolution unless the exact energy losses can be known for each bin. There is also a loss of resolution in Figure 6.39 in using the true losses from $P_{10m}(b)(dE_{MMC}; E)$ to $P_{100m}(c)(dE_{MMC}; E)$ due to the larger segment length as in Section 5.2.1.

6.6 Selection Criteria for Improved Resolution

In this section we introduce two event selection criteria which result in an improved resolution which surpasses TE for selected events. The first uses the shape of the likelihood space to select well-reconstructed events, the second uses the path length of the track to select events.

6.6.1 Likelihood Selection

The likelihood selection approach was explored in order to be able to identify cases where the Edepillim reconstruction method was not working.

The Edepillim method reconstructs the initial muon energy by finding a minimum in the likelihood space. As explained in Section 6.5, because of the physical limitations it is possible for the likelihood space to become ‘stuck’ where the minimum corresponds to the sum of the energy losses i.e. $E_0 = \sum dE$. In cases where the likelihood space is ‘stuck’ the energy reconstruction would be defined as not working as no reconstruction based on probabilities of energy losses is being made and thus it is not a reconstructed energy which should be trusted, except in cases of where the muon runs out of energy along its track.

The “likelihood selection” checks for reconstructed energies that are within a limit of

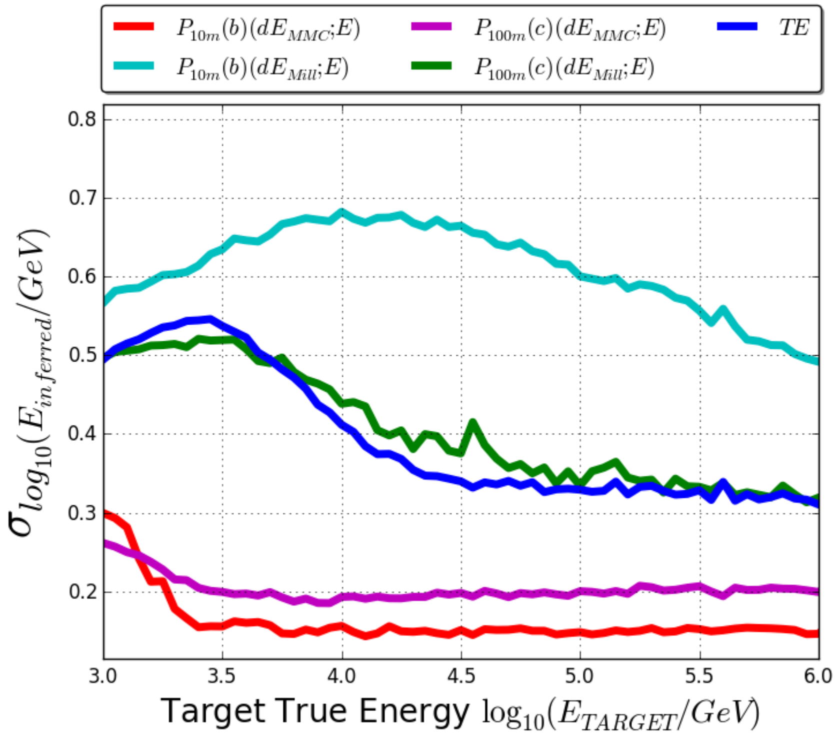


Figure 6.39: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) and true losses (MMC) with TE shown for comparison. The improvement of resolution from previous reconstruction $P_{10m}(b)(dE_{Mill}; E)$ to $P_{100m}(c)(dE_{Mill}; E)$ where the reconstructed energy losses are closer estimates to true energy losses and the threshold losses do not contribute. The segment length $X = 100\,\mathrm{m}$ is used as it was a consistent resolution across target energies. The resolution for Edepillim using Millipede losses (*green*) is now approaching that of TE (*blue*) but is still not as good as the case of using the true losses (*pink*).

the sum of the energy losses such that the selection cut is given by

$$E_0 > \text{limit} \times \sum d\vec{E}_{Mill} \tag{6.1}$$

The events below this cut are not trusted reconstructions and can be removed from the resolution distribution.

The selection cut is done on Edepillim reconstructed events where the energy loss pattern accounts for continuous energy losses and only has losses contained in the detector volume. The value for the limit of the selection cut can be varied as shown in Figure 6.40, with the limits of 2 and 4 resulting the same reconstruction resolution, and the higher limits of 8 and 16 resulting in worse resolutions. The higher the limit the more events will be removed using the cut, with limits 2, 4, 8 and 16 keeping 80%, 56%, 22% and 7% respectively. The selection limit of 2 will be taken for comparison as it has the best reconstruction resolution with the largest number of muon events kept.

The comparison of events kept to events removed is shown in Figures 6.41 and 6.42, where Figure 6.42 shows the Edepillim reconstruction for all the events removed. Some of the events removed were accurately reconstructed and form a diagonal peak line, however many were not and the cause of the spread of the distribution with a large width in the reconstructed energy.

In Figure 6.43 a comparison of TE is made to the Edepillim reconstruction resolution for muon events which have passed the likelihood selection with a limit of 2. The reconstruction results for Edepillim before cuts are also shown. The optimal rebinning for these selected events was found to be $X = 100\,\mathrm{m}$ with the resolution for the selection cut $P_{100m}(d)(dE_{Mill}; E)$ approaching the TE resolution for the same events.

This selection method, keeping events with $E_0 > 2 \times \sum d\vec{E}_{Mill}$, creates a cut on the $X = 100\,\mathrm{m}$ rebinned events with approximately 80% of events kept. This cut results in Edepillim performing with the same resolution as TE. However the method does demonstrate an additional power of Edepillim in that there is a way to tell if the reconstruction is working, by checking for a “stuck” likelihood. In the next section, we show how Edepillim does outperform truncated energy on events with the longest path lengths through the detector.

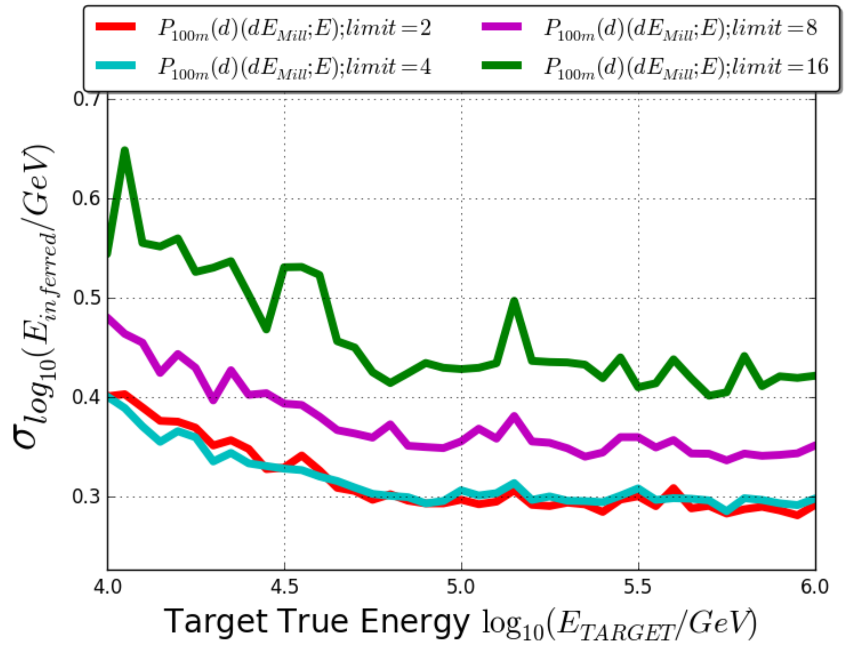


Figure 6.40: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with a likelihood selection of events with limit 2, 4, 8 nd 16. The likelihood selection cut has the best resolution for limit 2 and 4 with the resolution becoming worse for the higher limits due to more well reconstructed events being removed with the cut. The higher limits also have more fluctuations in the resolution due to the reduced number of events. The selection limit of 2 will be used as the optimal cut as it has good resolution as well as keeping the largest number of events.

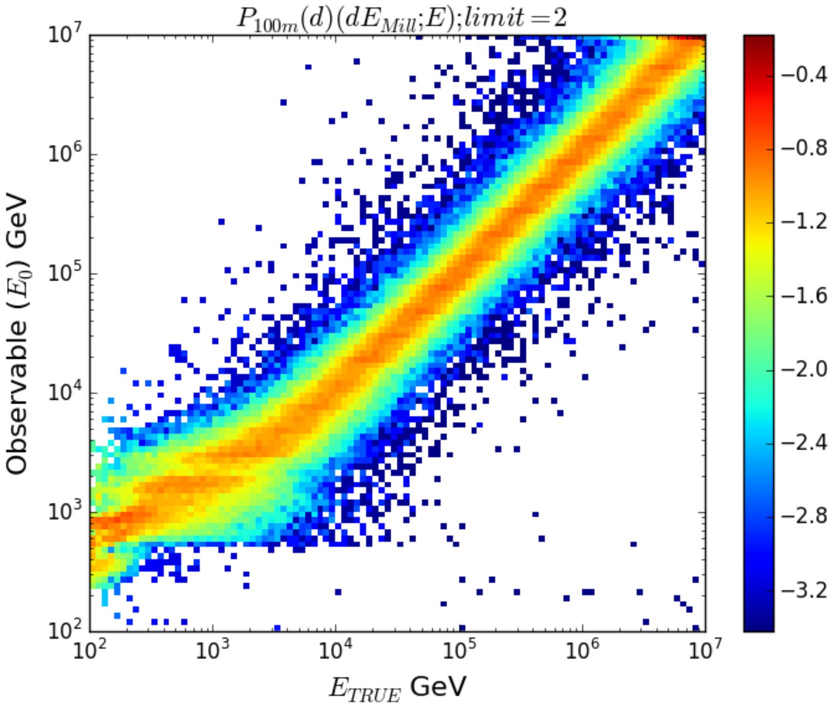


Figure 6.41: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy losses pattern (Mill) that has had segment length rebinned to $X = 100\,\mathrm{m}$. A selection cut has been applied on events following Eq 6.1 with a limit of 2 resulting in 80% of events remaining in the distribution. The cut has removed outlier events from the distribution which will improve the reconstruction resolution.

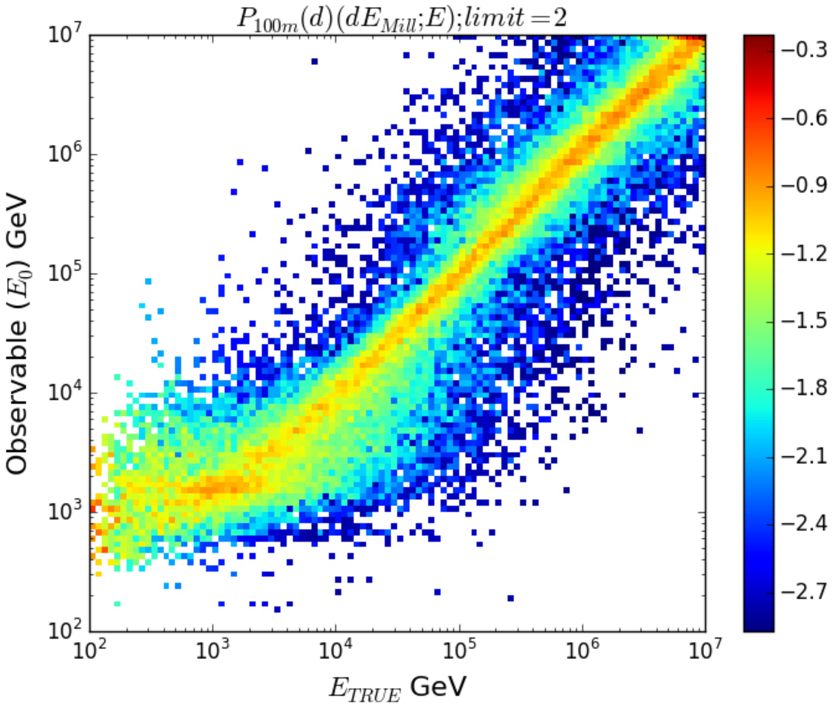


Figure 6.42: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy losses pattern (Mill) that has had segment length rebinned to $X = 100\,\text{m}$. These events did not pass the selection following Eq 6.1 with a limit of 2, thus these remaining 20% of events have been determined to be stuck likelihoods. There is a larger width to these events, whose removal has improved the reconstruction distribution in Figure 6.41.

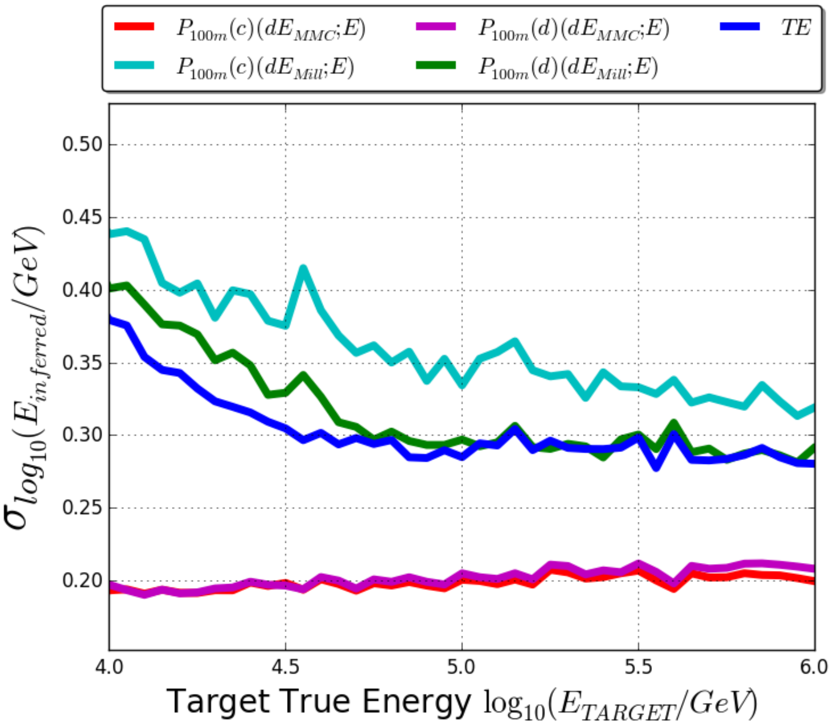


Figure 6.43: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with a likelihood selection of events. There are 80% of events kept in this selection $P_{100m}(d)(dE_{Mill}; E)$, only the same events for Edepillim on true losses and TE are shown. Also shown for comparison is the resolution of the full set of events the Edepillim reconstruction with rebinned segment length with threshold energy losses removed from the likelihood calculation $P_{100m}(c)(dE; E)$ from Section 6.5.2 for reconstruction on both Millipede and true energy losses. The likelihood selection cut has improved the Edepillim reconstruction to the resolution of TE.

6.6.2 Path Length Selection

In order to look for improvement of the resolution, the effect of the total path lengths of the events was explored. In Section 5.2.2 it was shown in Figure 5.23 that for the case of the true energy losses the length of muon track has an effect on the level of resolution. For the IceCube simulation in this chapter a path length cut selection was made at various path lengths and the resulting resolutions are shown in Figure 6.44, with a cut on events of $L > 1000\,\mathrm{m}$ showing the greatest improvement in resolution. This cut resulted in only 45% of the simulation events remaining in the reconstruction distribution. The reconstruction distibution for the optimal path length cut is shown in Figure

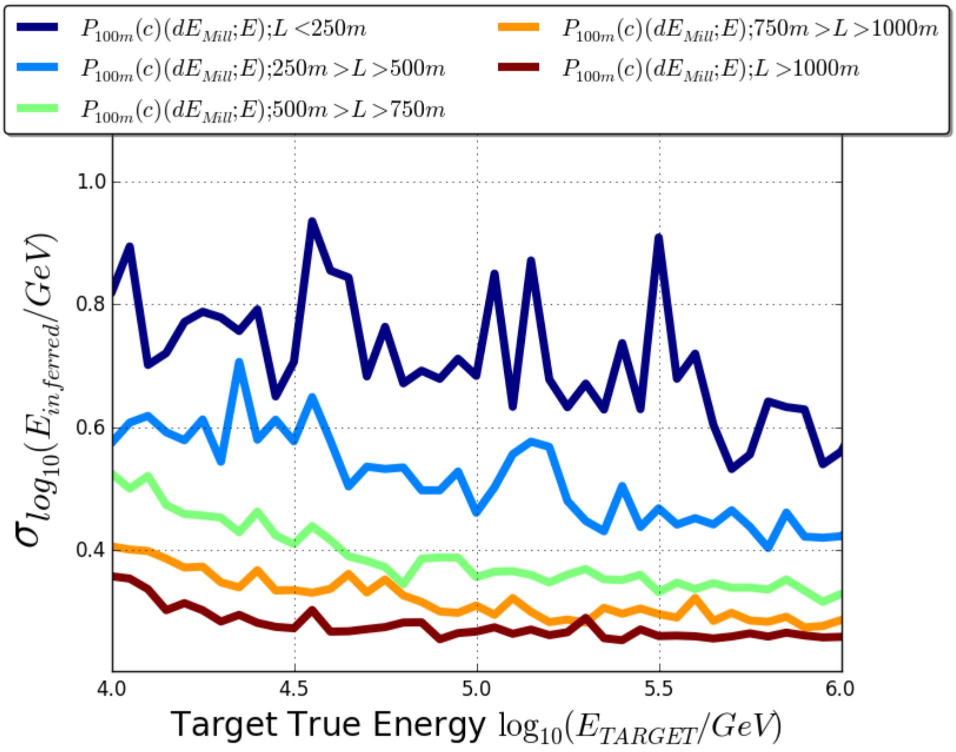


Figure 6.44: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with various path length cuts. All events have had the contribution of threshold energy losses removed from the reconstruction and only events within the range path length L are included. The path length cut of $L > 1000\,\mathrm{m}$ provides the best resolution for the events.

By performing the cut on the path length of events the resolution of the remaining events improves, relative to previous methods as shown in Figure 6.46. Segment lengths

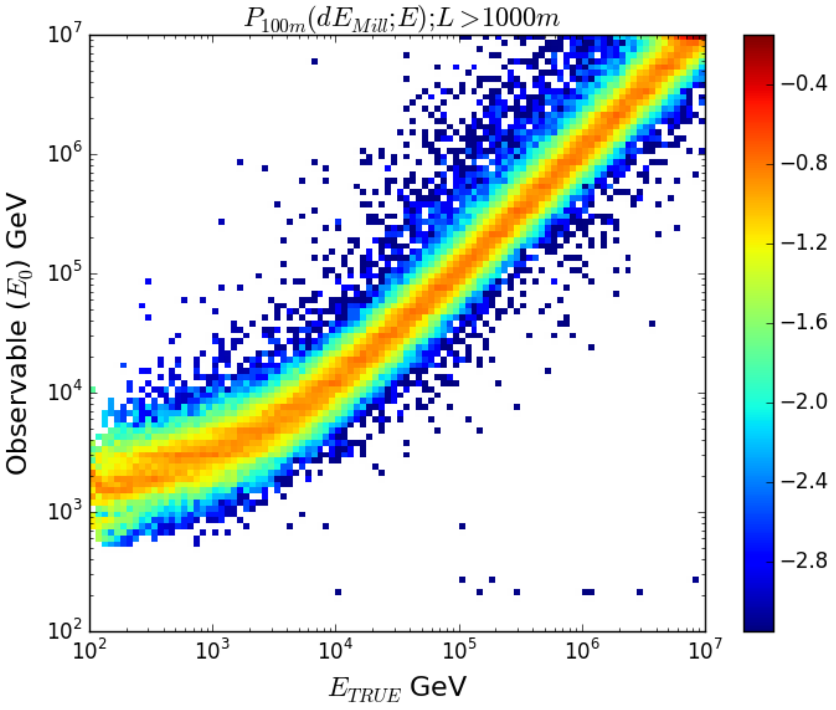


Figure 6.45: Reconstruction distributions of the observable muon reconstructed energy E_0 vs the true muon energy E_{TRUE} for a reconstructed muon energy losses pattern (Mill) that use segment length $X = 100\,\mathrm{m}$ and all threshold losses are ignored in the likelihood. Only the events with path length $L > 1000\,\mathrm{m}$ are included in the distribution. This reconstruction distribution results in the best resolution for Edepillim reconstruction on Millipede losses.

$X = 60 - 200\,\mathrm{m}$ now have resolution between $\sigma = 0.25 - 0.35$ compared to Figure 6.37 where the same segment lengths were between $\sigma = 0.3 - 0.4$.

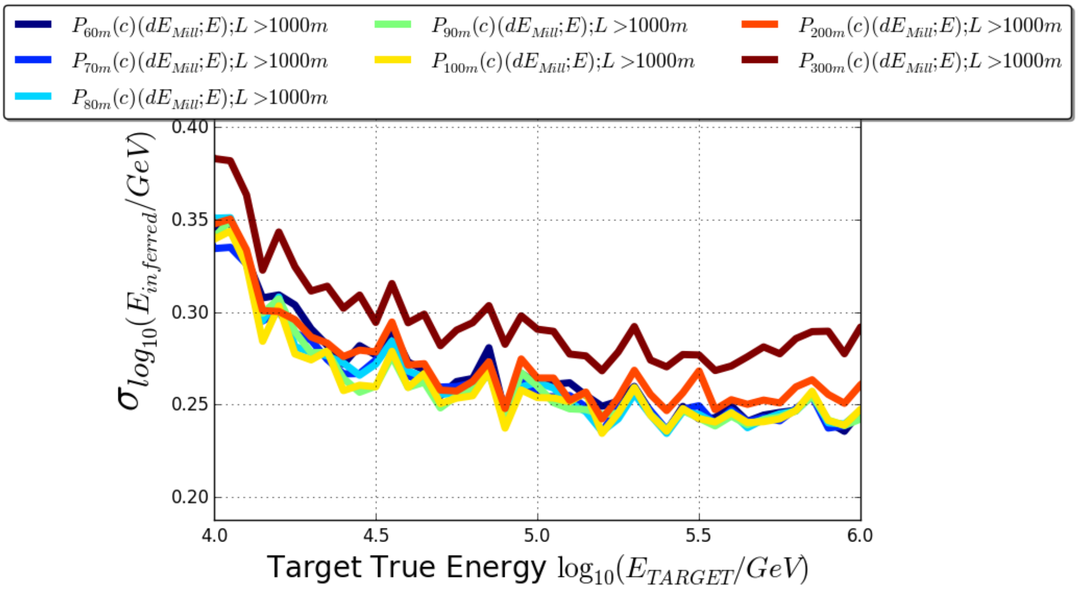


Figure 6.46: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) with various rebinned segment lengths. All events have had the contribution of threshold energy losses removed from the reconstruction and only events with path length $L > 1000\,\mathrm{m}$ are included. As in Figure 6.37 there is no clear optimal segment length across all target true energies, so the reconstruction of $X = 100\,\mathrm{m}$ will be used.

Given that the segment length of $X = 100\,\mathrm{m}$ has a consistent resolution performance across all energies, a comparison is made to the true losses and TE method in Figure 6.47, with the Edepillim reconstruction on Millipede losses $(P_{100m}(c)(dE_{Mill}; E); L > 1000m)$ $(green)$ having a much better resolution than the reconstruction for all events $P_{100m}(c)(dE_{Mill}; E)$ $(cyan)$. Of important note is that with the path length cut the Edepillim reconstruction is now out-performing the TE reconstruction $(blue)$ for the same events.

In all muon energy reconstructions a cut on path length will improve the reconstruction resolution, as both TE and Edepillim perform better on longer muon path lengths. This is shown in Figure 6.48 where the reconstruction resolution for short events with path length $L < 1000\,\mathrm{m}$ can be compared against events with path length $L > 1000\,\mathrm{m}$. The

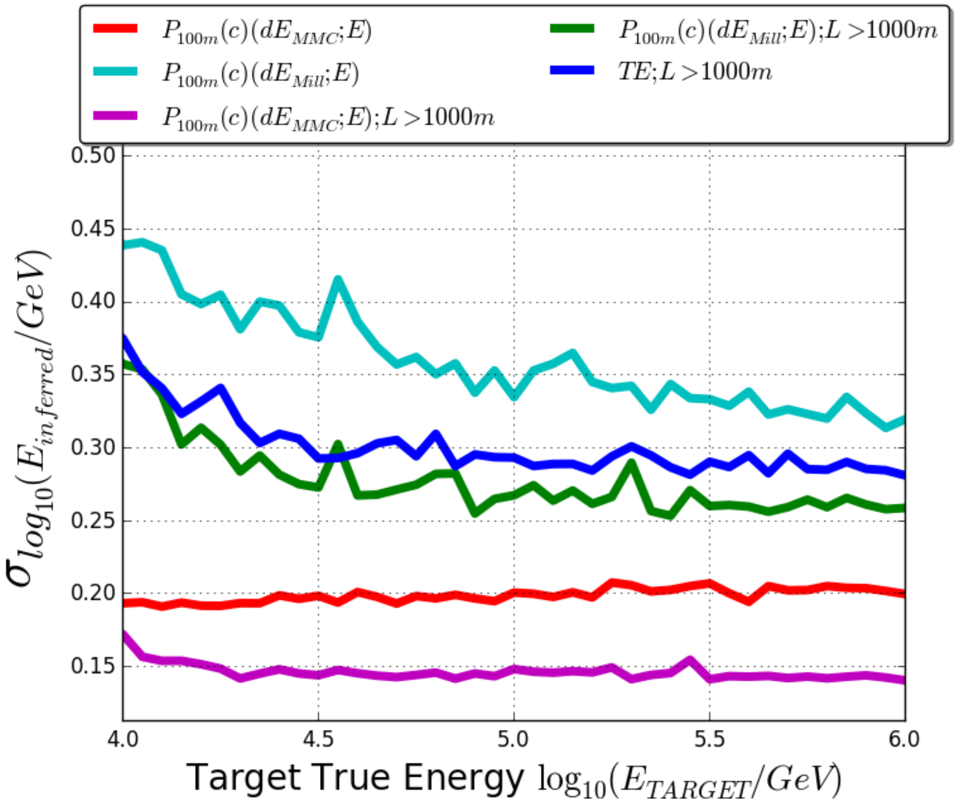


Figure 6.47: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) and the true losses (MMC) with the same TE events for comparison. All events have had the contribution of threshold energy losses removed from the reconstruction and only events with path length $L > 1000\,\mathrm{m}$ are included. The Edepillim energy reconstruction on Millipede (*green*) now has a better resolution then the TE (*blue*).

short events (*red*, *yellow* and *cyan*) all have a worse reconstruction resolution than the long events (*pink*, *green* and *blue*). This is an effect on all energy reconstructions as the amount of information is reduced in shorter path lengths, and the reconstructions are less able to perform with a good resolution.

However the improvement in resolution for the path length cut is different for different reconstructions and can be observed as the gap between the long and short reconstructed events in Figure 6.48. For TE the gap between the long (*blue*) and the short (*cyan*) reconstructed events is approximately on average a loss of the resolution measure of $\Delta\sigma = 0.05$. This can be compared to the Edepillim reconstruction using Millipede reconstructed energy losses, with the gap between the long (*green*) and short (*yellow*) events being on average over all energies $\Delta\sigma = 0.1$, this is quite a large improvement in resolution. The same gap in resolution between long and short events is present in the Edepillim reconstruction on true energy losses (short:*red*, long:*pink*) where the change in resolution measure is approximately $\Delta\sigma = 0.1$, which is the same as the change in resolution for Edepillim reconstruction on Millipede energy losses. This indicates that the loss in resolution from path length cuts is fundamental to all Edepillim reconstruction performance, confirming what was shown in Chapter 5.

This gap in Edepillim resolutions for long and short events can account for the loss in the overall Edepillim reconstruction resolution relative to TE when assessed on all events. Thus by removing short path length events the Edepillim resolution performs consistently better than TE for the same long events. This shows that even without improvement to the reconstructed energy losses to the level of the true energy losses, Edepillim can still achieve a better reconstruction resolution than TE, for long path length events.

6.7 Alternative Production of PDF using Millipede Reconstructed Losses

At the beginning of its development Edepillim reconstruction did not use PROPOSAL to produce PDFs but instead used the Millipede reconstructed energy losses of simulated events. This was found to not work with a good reconstruction and the reconstruction method moved to using true losses in the PDF. In this chapter we have outlined the pro-

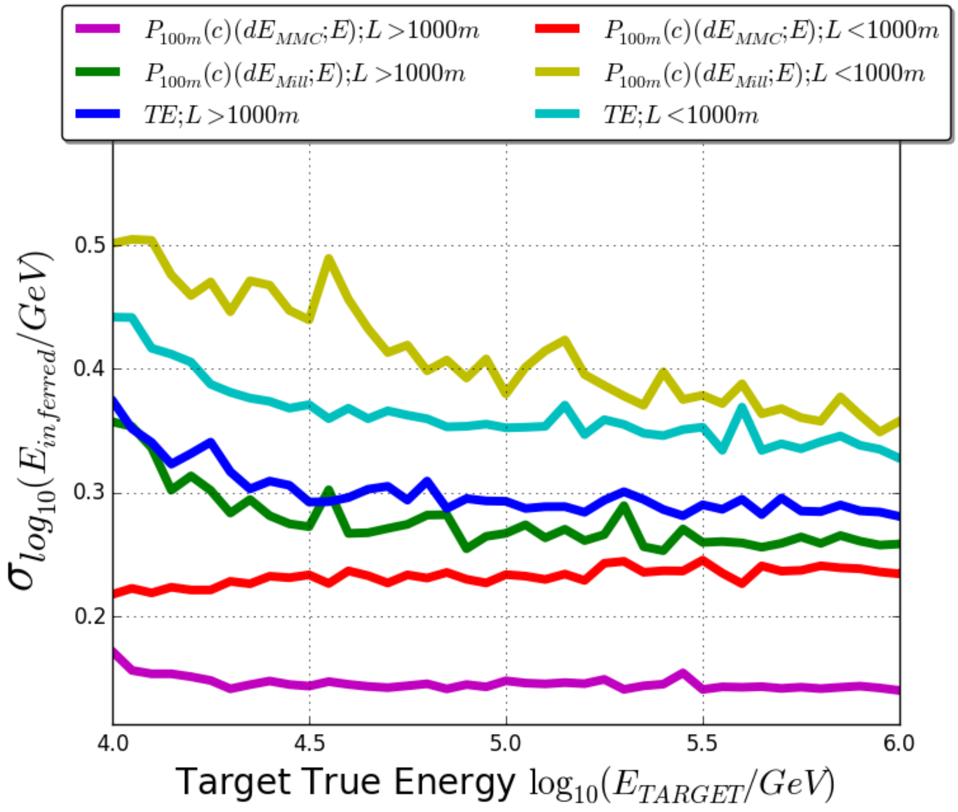


Figure 6.48: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on Millipede losses (Mill) and the true losses (MMC) with the same TE events for comparison. All events have had the contribution of threshold energy losses removed from the reconstruction and with path length $L > 1000\,\mathrm{m}$ or $L < 1000\,\mathrm{m}$. The large gap between the Edepillim resolution for short and long path lengths explains the loss of resolution when all events are used.

cedure to make the Millipede reconstructed losses appear more like the true losses so that the reconstruction can perform at its best. An alternative to that is to go back to starting with a PDF that describes not what the most likely true loss was, but what the most likely Millipede reconstructed loss was. In this alternative method the PDF is built from the Millipede reconstructed losses dE given their true muon energy $P(Mill)(dE;E)$. The energy losses come from the reconstructed losses in the simulation. In this construction the error in the reconstructed losses is contained within the PDF, as shown in Figure 6.49, in this case the continuous losses are still being added to the segment length so there is a large number of losses at $dE = 20\,\text{GeV}$. For testing purposes only half of the simulations muon's energy losses are used to build the PDF, the rest of the muon events will be used for testing the reconstruction. This avoids overlap in the events entered into the PDF and in the reconstructed events to prevent possible bias.

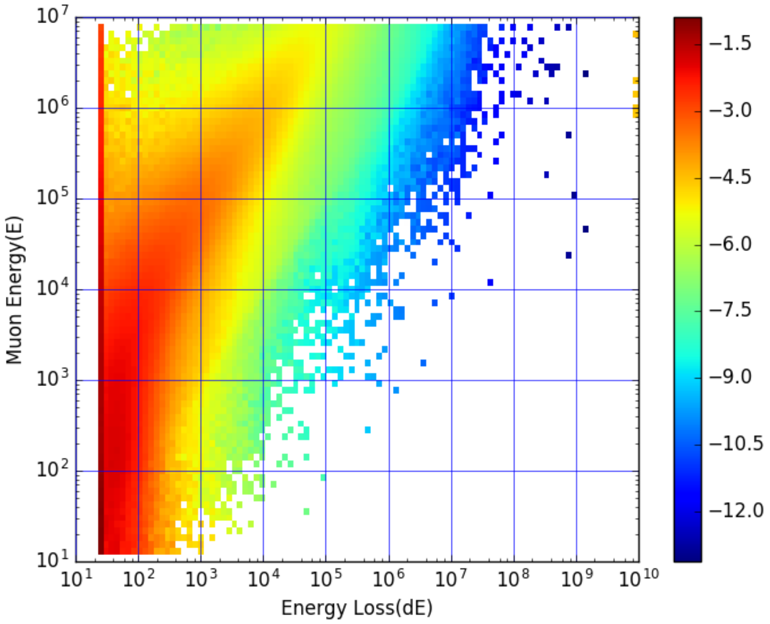


Figure 6.49: Probability distribution function built from Millipede reconstructed losses of rebinned segment length $X = 100\,\text{m}$ binned along energy loss dE. The muon energy is taken as the true energy before each segment loss. There is a large probability at the threshold energy loss due to the abundance of zero losses in the Millipede reconstruction that have the continuous energy loss added.

In the case of the reconstruction resolution shown in Figure 6.50 the Millipede PDF was used with a rebinned segment length of $X = 100\,\mathrm{m}$, and only events with path length $L > 1000\,\mathrm{m}$ were used. Note that continuous losses were still accounted for in both the energy loss pattern and the PDF, but no threshold losses were removed. The comparison in Figure 6.50 shows that there is no significant difference between using the Millipede PDF and using the procedure outlined in this chapter using a PROPOSAL generated PDF. This means the procedure outlined in this chapter to improve the muon reconstruction allows for the errors in the Millipede energy losses to be accounted for in both the PDF and energy loss pattern.

This method of using the Millipede PDF would require the PDF be built using the same energy loss reconstruction settings as in the required simulation. It is possible that the inaccurate Millipede energy losses are introducing an additional error to the Edepillim reconstruction in that the muon energy at any given point E_N is given by the test muon energy and the sum of the previous losses, if the losses are inaccurate this muon energy will also be, thus it may be looking at the wrong probability in the PDF. However, there was no difference found in the reconstruction result if the PDF is built using the true muon energy along the track in the PDF $P(dE; E)$, or the muon energy based on previous Millipede losses i.e. $E_N = E_0 - \sum dE_{Mill}$. This indicates that the error from assuming the muon energy along the track based on Millipede energy losses is negligible, especially compared to the error on the energy loss itself.

The reason for this alternative PDF working now when it did not in its initial use is that other factors need to be accounted for in order to achieve good reconstructed resolution, the biggest being that segment losses still need to be rebinned to achieve the best resolution. If the reconstructed Millipede losses are rebinned such that they match the true losses, then building the PDF, using either method should lead to essentially the same result, as we have seen. Also continuous losses still need to be included in the energy loss pattern to be properly accounted for in the PDF.

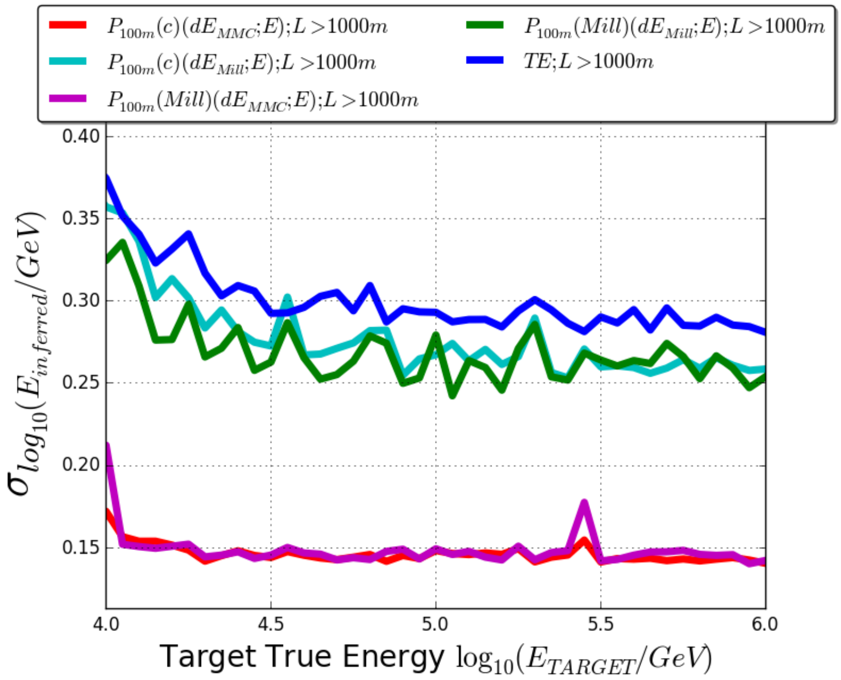


Figure 6.50: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method using a PDF built from Millipede reconstructed energy losses on Millipede losses (Mill) (*green*) and the true losses (MMC) (*pink*) with the same TE events for comparison. The reconstruction are compared to the previous results where events have had the contribution of threshold energy losses removed from reconstruction. For all resolutions only events with path length $L > 1000\,\text{m}$ are included. The reconstitution using a Millipede PDF (`green`) is working at the same resolution as the previous results (*cyan*) indicating all the Millipede errors have been accounted for by the previous result.

6.8 Future Improvements of Millipede Energy Losses

An alternative approach to overcoming the error on the energy losses would be an energy loss reconstruction with a combined likelihood fitter which incorporates the knowledge Edepillim has of the most likely energy loss with Millipede's reconstruction of the event seen. This would be done by reconstructing the muon event using Millipede and using the likelihood output to combine with the likelihood of Edepillim for those losses via

$$-\log L_{joint} = -logL_{Millipede} - \lambda \log L_{Edepillim} \tag{6.2}$$

where λ is a weighting factor which will allow each likelihood equal contribution such that one likelihood would not overwhelm the other. The likelihood equation would be used to perform a joint minimisation to find the energy losses that explain simultaneously what is seen by Millipede and what is expected by Edepillim[1]. The hope is that the joint likelihood would reduce the large unphysical energy losses reconstructed by Millipede, as they are not consistent with the expectation for losses along the path length.

6.9 Summary

In this chapter we have tested the Edepillim muon reconstruction method on IceCube simulations using the reconstructed energy losses from Millipede. We showed that initially the Edepillim reconstruction does not work with good reconstruction resolution, due to problems with the Millipede reconstructed energy losses, where the reconstructed losses did not match the true losses, and also, that there was an abundance of zero losses in the energy loss pattern. The zero losses were not being correctly accounted for in the PDF because of limits on the energy loss range. To provide consistency the continuous losses were accounted for in the energy loss pattern and the PDF.

To remove the incorrectly reconstructed large energy losses that are typically placed at the edges of the energy loss pattern a cut was performed on the energy loss pattern

---

[1] Development of this method has begun but has not been completed to a level to be included in this work.

to only include losses that occur within the IceCube detector volume. This reduced a number of large mis-reconstructed losses leading to better consistency between the true total energy loss and the total reconstructed energy loss.

In another section we demonstrated how too many zero or threshold losses can cause the Edepillim reconstruction to simply return the sum of all the energy losses. This is not a working Edepillim reconstruction with no minimum in the likelihood being found and would only be correct in the cases where the muon runs out of energy. Since so many Millipede reconstructed events have a large fraction of threshold losses, this effect is dominating the Edepillim reconstruction causing a poor reconstruction resolution.

In order to improve on the accuracy of the reconstructed energy losses the energy loss pattern was rebinned to larger segment lengths. This increased the accuracy of all reconstructed energy losses with respect to the true energy losses and allowed for a better reconstruction resolution. The threshold losses can also be removed from contributing to the Edepillim likelihood to further improve the reconstruction to the level where the resolution is the same as that of TE.

Two selection criteria were tested on the reconstructed events. The first performed a likelihood selection based on the reconstructed value within a limit of the sum of the energy losses. This selection allowed threshold losses to still contribute to the likelihood and shows that Edepillim has the ability with its likelihood to assess that events have been well reconstructed. The selection criterion with the best results was a cut on the muon path length, when only events that had a long path length of greater than 1000 m are included in the reconstruction. Here Edepillim has a better resolution than TE on the same events. This improvement with path length is consistent with the results in Chapter 5 which found that Edepillim works best with longer path lengths. Thus the Edepillim reconstruction has a better resolution than TE in the case of long events.

Finally a comparison was done for the final result of Edepillim using a path length cut, with an alternative method of using a PDF built from Millipede energy losses in simulation and performing an Edepillim reconstruction. This method uses the PDF to account for the errors in Millipede, and has the same resolution as the main results of the Chapter, when it also accounts for continuous losses and rebinned segment length. This method is now working because it is rebinning the segment lengths such that the

Millipede losses are closer to the true losses.

The work in this chapter shows that with the procedures outlined, for improving the Millipede energy losses and including at cut on the events path length, the Edepillim reconstruction has a better reconstruction resolution than TE. There is also potential for more improvement to the resolution of Edepillim on the true losses, if the reconstructed losses could be initially reconstructed to the true losses by Millipede.

Chapter 7

Assessing the Energy of IceCube's Highest Energy Event

IceCube is capable of detecting muon track events from high energy astrophysical neutrinos, with the highest detected muon event to date having PeV-scale deposited energy. The Edepillim method has the potential to perform best for such high energy events with large stochastic losses more easily being reconstructed in the energy loss pattern. In this chapter we discuss Edepillim's performance on simulation events similar to the highest energy detected event and show the reconstructed energy for the event using Edepillim compared with the reconstructed energy for truncated energy (TE).

7.1 Multi-PeV Event Description

On June 11th 2014 IceCube observed the event with the highest deposited energy to date. It had 2.6 PeV of energy deposited in the detector, shown in Figure 7.1. The event has a high likelihood of being astrophysical in origin as the p-value of this event being atmospheric is less than 0.005% [21].

In its first publication [21] the event was reconstructed using the standard energy reconstruction of truncated energy using the AllDOMs method. The muon energy was calculated as 4.5 ± 1.5 PeV. The event's posterior probability distribution in neutrino energy is shown in Figure 7.2, with the three curves showing the probability for the three

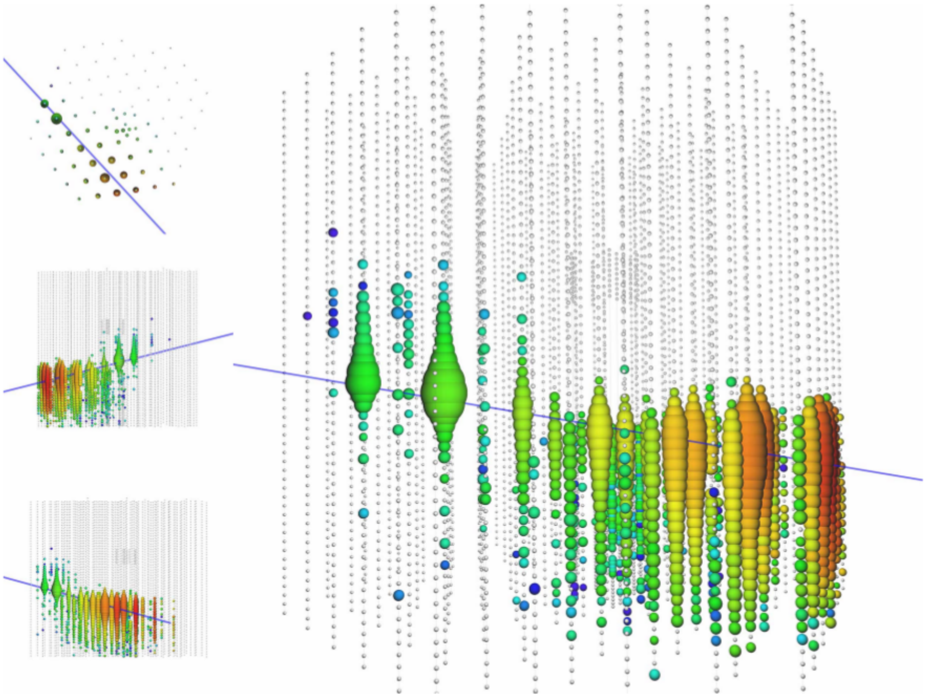


Figure 7.1: Event view of the highest energy event ever seen by IceCube with 2.6 PeV of deposited energy in the detector.

possible neutrino flavours; muon, tau and electron. The total probabilities for each flavour are 87.7% for muon neutrino, 10.9% for tau neutrino and 1.4% for an electron neutrino. Thus the event is most likely to be a muon neutrino, which would have a median neutrino energy of 8.7 PeV.

The multi-PeV event is an optimal example of the potential for using the Edepillim reconstruction method. The energy loss pattern of the reconstructed losses shown in Figure 7.3 show that it is a high energy event with a high fraction of radiative stochastic losses. Additionally it is a muon track that travels through the detector, with a large path length for its energy loss pattern of 1150 m. These characteristics lead to improvement in the effectiveness of the Edepillim energy reconstruction. In this chapter we reconstruct the event's energy using Edepillim and compare against simulations of similar events.

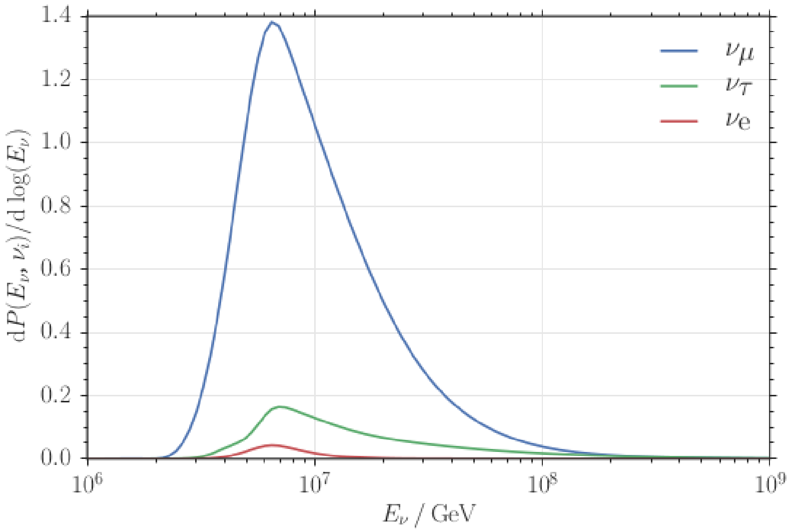


Figure 7.2: The probability distribution function for neutrino energy for the multi-PeV event. Each curve represents a different neutrino flavour, ν_μ, ν_τ, ν_e.

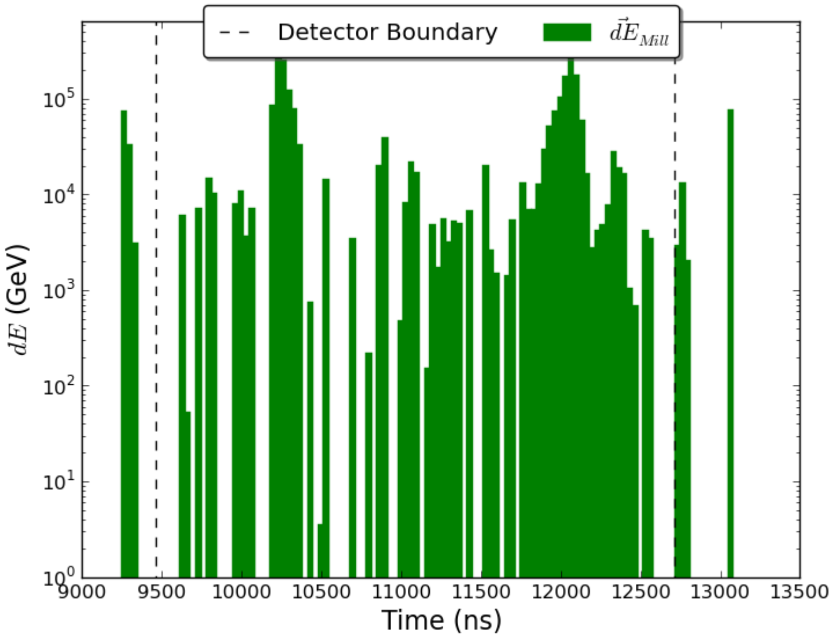


Figure 7.3: The Millipede reconstructed energy loss pattern for the multi-PeV event with segment length $X = 10\,\mathrm{m}$. The dashed line show the detector boundary for the event's track. It is a long track with most of its energy contained in the detector volume and a large fraction of stochastic losses.

7.2 Simulation of Similar Multi-PeV Events

In order to test the effectiveness of Edepillim on such an event a simulation of 50,000 similar events was used. The events varied around the original event best fit direction by 1 degree, with energies in the range $10^5 - 10^8$ GeV with equal numbers of muon events per log energy bin.

As all the events are simulated with a similar trajectory all the events had the same path length of 1160 m and so no cuts had to be made on path length. All the events are high energy events with a large amount of stochastic losses. In Figure 7.4 a simulated muon event's energy loss pattern is shown. In Figure 7.5 the same event has had its segment length rebinned to 90 m and its reconstructed energy loss pattern has no bins with zero energy loss bins, and its loss pattern more closely resembles that of the true energy loss pattern.

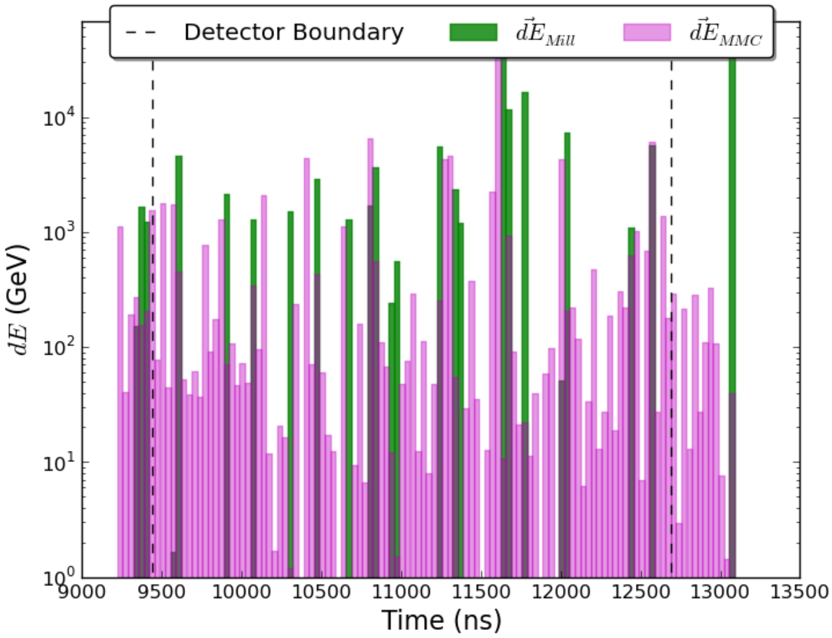


Figure 7.4: A simulated muon's energy loss pattern for its true losses $d\vec{E}_{MMC}$ and its reconstructed losses $d\vec{E}_{Mill}$, for an event with energy $E_{TRUE} = 3.7 \times 10^5$ GeV and path length $L = 1160$ m, with the same trajectory through the detector as the multi-PeV event. The dashed line shows the detector boundary for the track.

The procedure for energy reconstruction differs from that in the previous chapters as

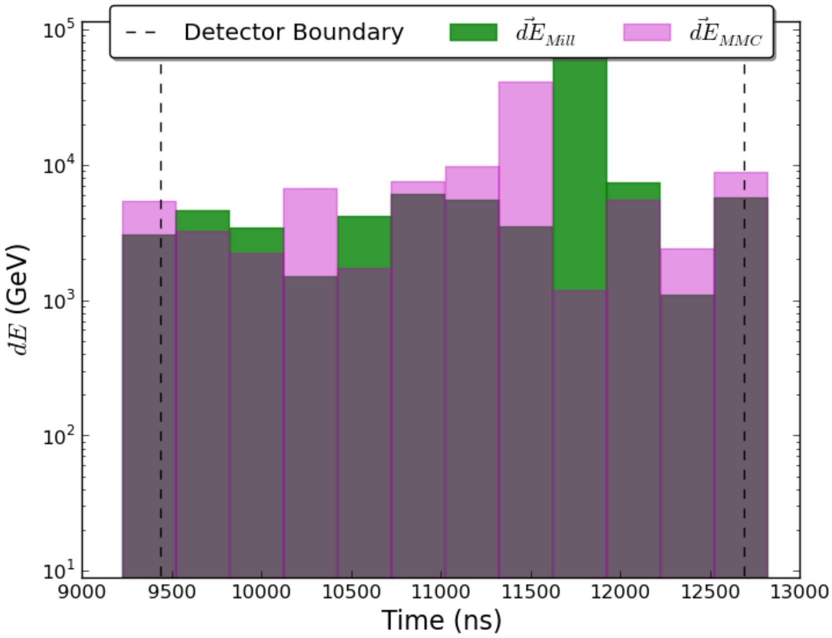


Figure 7.5: The simulated event from Figure 7.4 energy loss pattern for its true losses $d\vec{E}_{MMC}$ and its reconstructed losses $d\vec{E}_{Mill}$, for an event with energy $E_{TRUE} = 3.7 \times 10^5\,\text{GeV}$ and path length $L = 1160\,\text{m}$. The energy losses have been rebinned to segment length $X = 90\,\text{m}$. The dashed line shows the detector boundary for the track.

the energy losses in this simulation did not have as many of zero energy loss bins in the energy loss pattern. The histogram in Figure 7.6 shows, for this simulation of multi-PeV like events, the range of the fraction of above-threshold losses in the $X = 10\,\mathrm{m}$ case. In this case 59% of the multi-PeV like muon events are below the 40% ratio threshold, where the event will have too many threshold losses for Edepillim to perform a working reconstruction, as explored in Section 6.5 for the case of simulation with lower energy muons.

The energy losses in this case were rebinned and once the segment length for the energy losses had been rebinned to $X = 90\,\mathrm{m}$, as shown in Figure 7.7, the percentage of muon events with too many threshold losses is negligible. We use $X = 90\,\mathrm{m}$ as in this work as we show later in this section based on simulation, that it produces the best reconstruction resolution, see Figure 7.8. This allows for the reconstruction to work without having to ignore the remaining threshold energy losses, as they no longer have any large impact on the energy loss pattern.

The reconstruction resolution for various rebinned segment lengths is shown in Figure 7.8. In these reconstructions no threshold losses needed to be removed, as in Chapter 6, as the rebinned segment length negated their influence. The optimal rebinned segment length with the best reconstructed resolution is $X = 90\,\mathrm{m}$ as it has the lowest value of σ across all target muon energies, though there is little variation between the rebinned segment lengths of $X = 60 - 100\,\mathrm{m}$.

The optimal rebinning of $X = 90\,\mathrm{m}$ is contrary to what was concluded in the previous chapter as being the optimal rebinning length, as in Chapter 6 it was harder to distinguish the resolution of the different curves using rebinned segment length. As this current simulation was a specific simulation tested for a very contained area of the detector it is not surprising that it yields a very clear but different result than a simulation that tries to take into account all directions through the whole detector. Additionally, the fractional energy loss implementation of Edepillim was used as it had a better reconstruction resolution for this case. This is due to the simulated events all having high energies above $10^5\,\mathrm{GeV}$, and the Edepillim reconstruction does not encounter the problems in the PDF in Figure 6.12 of the threshold losses on the diagonal line, as the likelihood rarely refers to the probabilities on the diagonal line. Thus the fractional implementation can be used.

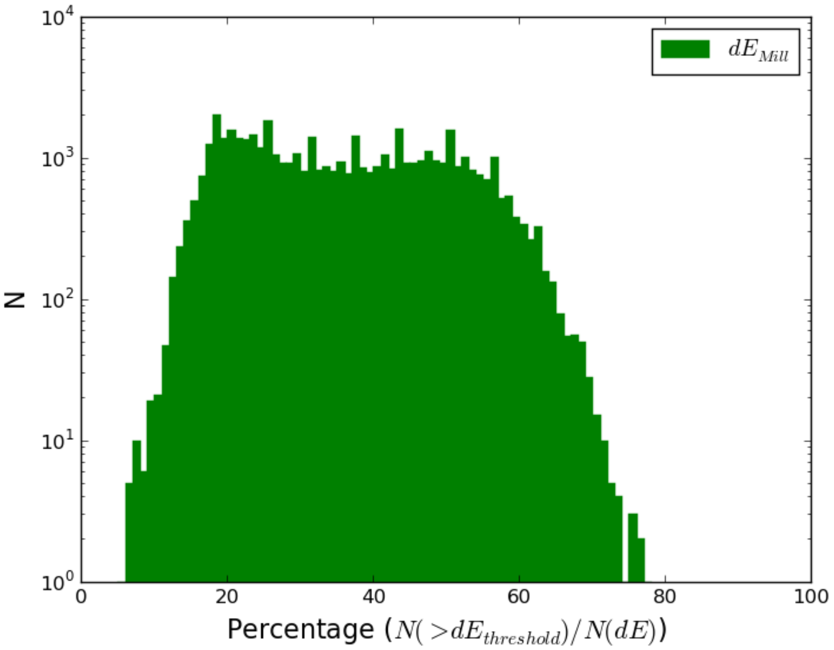


Figure 7.6: Histogram of the percentage ratio of above threshold losses in each muon event's energy loss pattern with segment length $X = 10\,\mathrm{m}$, for the simulation with energy range $10^5 - 10^8\,\mathrm{GeV}$. Events that have a ratio of above threshold losses below 40% will have too many zero losses in the pattern for Edepillim to be able to reconstruct the muon energy.

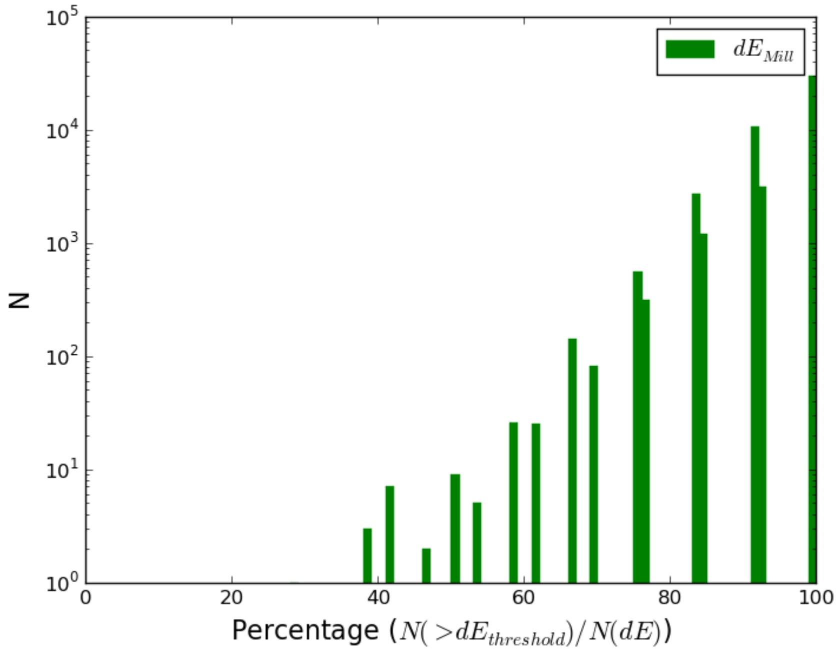


Figure 7.7: Histogram of the percentage ratio of above threshold losses in each muon events energy loss pattern with rebinned segment length $X = 90\,\mathrm{m}$, for the simulation with energy range $10^5 - 10^8$ GeV. Events that have a ratio of above threshold losses below 40% will have too many zero losses in the pattern for Edepillim to be able to reconstruct the muon energy. The binning effects in the histogram is due to the rebinned segment lengths causing only certain percentages to be filled as there is a lower number of energy loss bins.

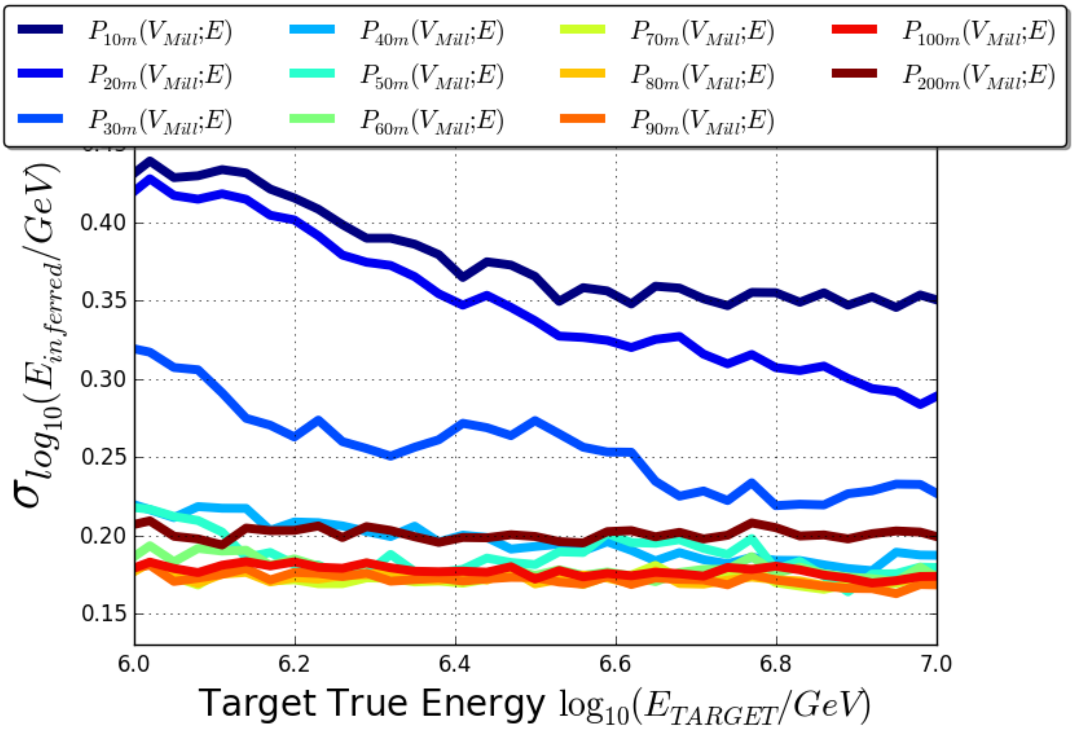


Figure 7.8: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the reconstructed losses (Mill) with various rebinned segment lengths X. In this reconstruction the fractional energy loss V method is used as it had the best resolution for the simulation.

In Figure 7.9 the reconstruction resolution for Edepillim is compared for the optimal rebinned segment length of $X = 90\,\mathrm{m}$ to TE and to using the true energy loss pattern. The results show a very clear difference in resolution for the Edepillim method compared to TE and in fact the resolution is very close to that of using the true energy losses. These results show the best performance of Edepillim on reconstructed losses in this work and indicate the potential of the method. If the simulation has a well reconstructed energy loss pattern the reconstructed resolution is close to the optimal resolution that you would get reconstructing on true energy losses.

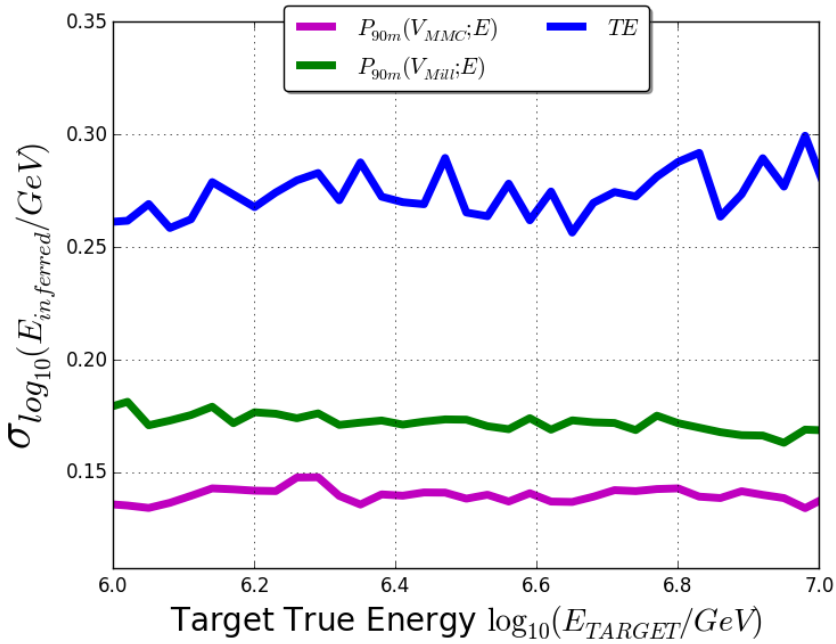


Figure 7.9: The ALF reconstruction resolution at target true energies, for the Edepillim reconstruction method on the reconstructed losses (Mill) and the true energy losses (MMC). The optimal rebinned segment length from Figure 7.8 is used $X = 90\,\mathrm{m}$. Shown for comparison is the TE resolution for the same events.

7.3 Energy Reconstruction for the Multi-PeV Event

In this section we use the Edepillim method to reconstruct the energy of the multi-PeV event using the optimal settings from Section 7.2. Using this reconstructed muon energy, in Section 7.3.2 we further reconstruct the neutrino energy of the event.

7.3.1 Edepillim Muon Energy Reconstruction

The energy reconstruction of the multi-PeV event was performed using the outlined procedure in the previous section with the preferred settings for optimal resolution as given by simulated events. The rebinned segment length of $X = 90\,\text{m}$ was shown to have the best Edepillim energy reconstruction resolution for events like the multi-PeV event. Figure 7.10 shows the Millipede energy loss pattern for this event using the rebinned segment length of $X = 90\,\text{m}$. The energy loss pattern shows that with rebinned segment losses there are no threshold energy bins which could affect the likelihood space, thus it is possible for the Edepillim reconstruction to work properly.

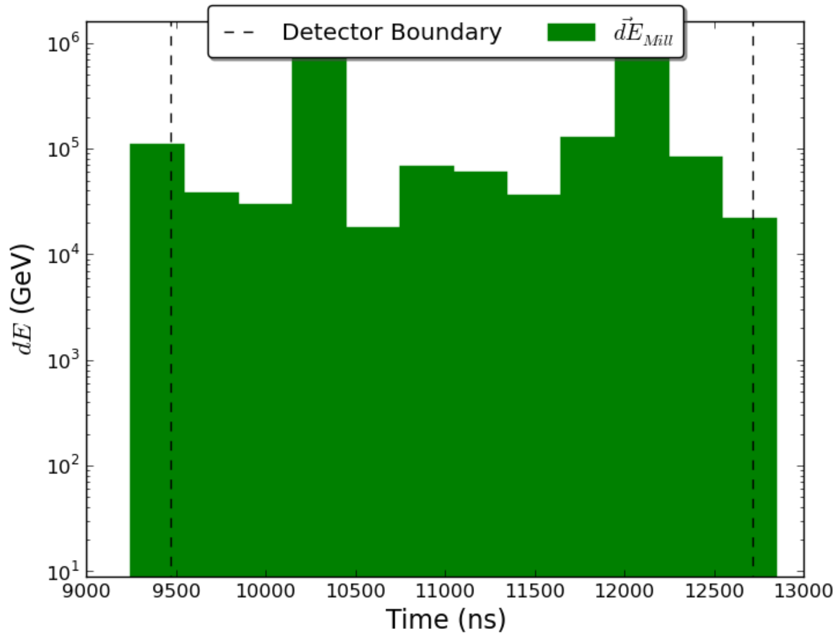


Figure 7.10: The Millipede reconstructed energy loss pattern for the multi-PeV event with energy losses rebinned from segment length $X = 10\,\text{m}$ to $X = 90\,\text{m}$. The rebinned segment length results in no threshold energy loss bins in the pattern which will allow for a good Edepillim reconstruction.

The Edepillim reconstruction's likelihood space for the multi-PeV event is shown in Figure 7.11. The Edepillim method reconstructs the event's energy to be $6.3^{+1.9}_{-1.1} \times 10^6$ GeV, which is a higher estimate than that of TE, $4.5 \pm 1.5 \times 10^6$ GeV. Also shown on the plot is the one-sigma confidence interval range on the reconstructed energy based on the Wilks theorem for the likelihood. While Edepillim reconstructs a higher energy for the event the two energies are consistent with each other as the maximum of TE, 6.0×10^6 GeV is greater than the lower limit of the confidence interval for Edepillim's reconstructed energy, which is 5.2×10^6 GeV.

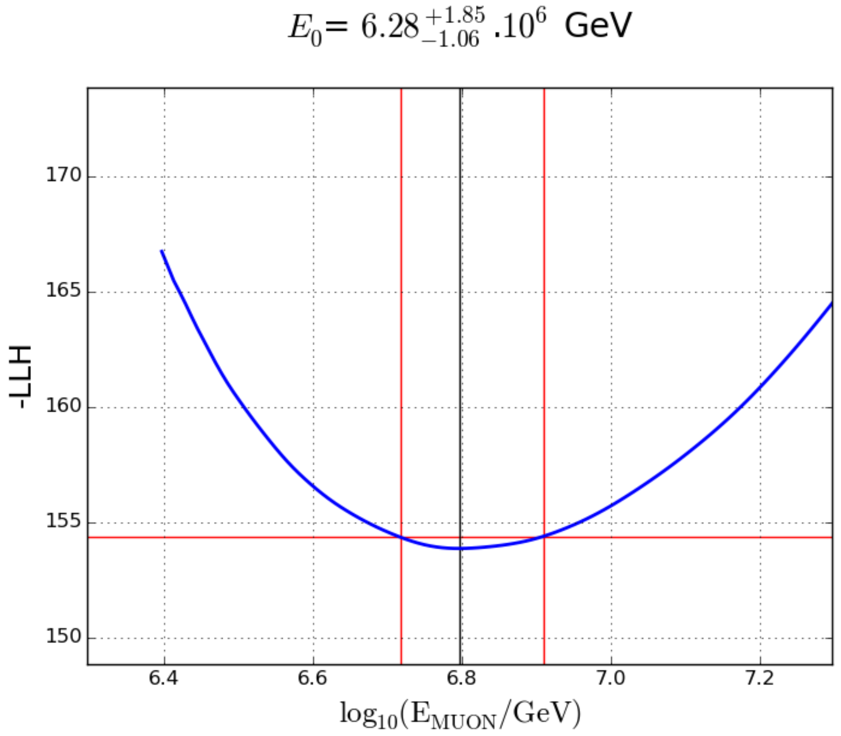


Figure 7.11: The Edepillim reconstruction likelihood space for the the multi-PeV event, for the case of segment length rebinning to $X = 90$ m. The likelihood space produces a smooth minimum, thus is a working Edepillim reconstruction. The black vertical line shows the minimum of the likelihood, and the vertical red lines are the range of the one-sigma confidence interval.

7.3.2 Neutrino Energy

The neutrino energy of an event can be estimated using the spectrum of the incident neutrino energies and the true muon energy as it enters the detector. This spectrum, as shown in Figure 7.12, describes all the possible neutrino energies a muon entering the detector could correspond to and is weighted to the best fit astrophysical flux found in [21]. Using the spectrum in Figure 7.12 and the Edepillim reconstruction distribution in Figure 7.13, a combined spectrum can be made by weighting the expected neutrino flux

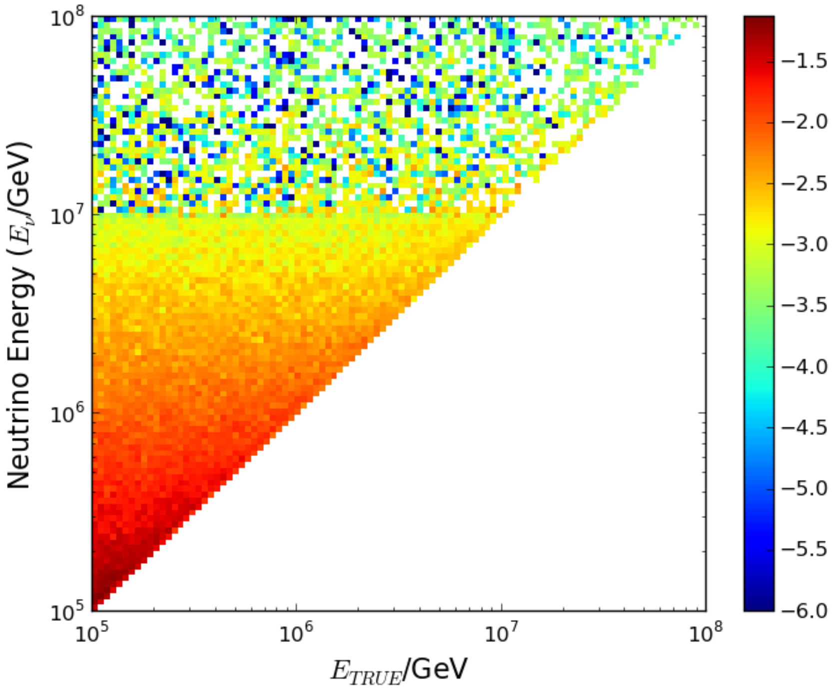


Figure 7.12: The spectrum of incident neutrino energy given muon true energy at the entry of the detector. The spectrum is weighted using the astrophysical weights found in [21]

.

for a muon true energy by the distribution for that true muon energy over reconstructed energies. This results in the combined spectrum in Figure 7.14 that now relates the neutrino energy to the reconstructed energy. The vertical lines that appear in Figure 7.14 are an effect of the reconstruction distribution in Figure 7.13 having reconstructed observable energies that are slightly less favoured than other energies. This effect cannot be easily seen in the reconstruction distribution but when the distributions are layered in Figure 7.14 they become clear. After extensive investigation the exact source is not well understood but the cause of these energies being disfavoured is likely to be due to the nature of the likelihood space, some specific reconstructed energies appear to be disfavoured, but here will not affect the rest of the work in this Chapter.

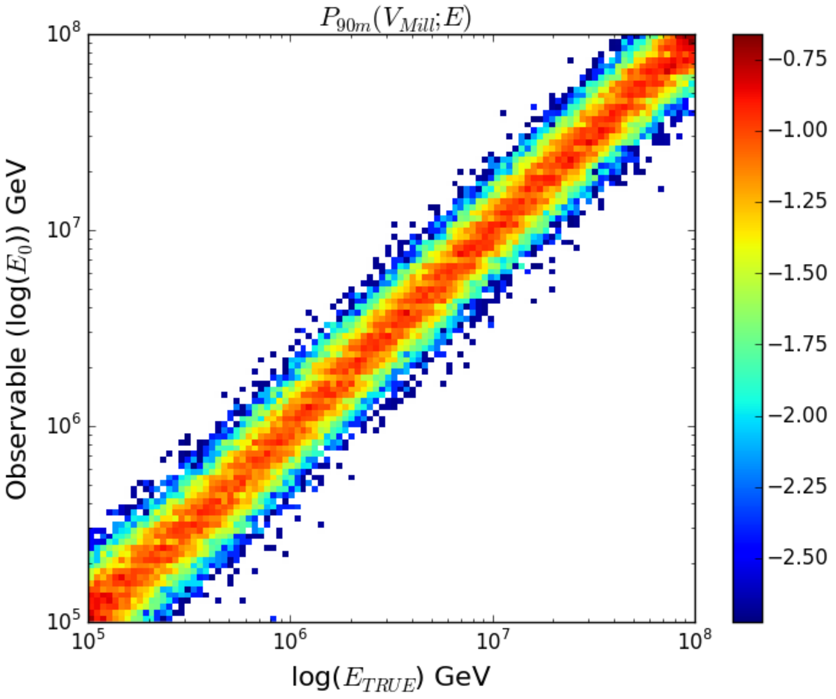


Figure 7.13: The Edepillim reconstruction distribution for an high energy muon events with segment length $X = 90\,\mathrm{m}$.

The neutrino energy spectrum for the multi-PeV event can be taken as a vertical slice in the spectrum in Figure 7.14 corresponding to the actual reconstructed muon energy $E_0 = 6.3 \times 10^6$ GeV. The resulting neutrino energy distribution as shown in Figure 7.15 can be compared to that in Figure 7.2 which is the event's neutrino energy spectrum found in [21] using TE. The median energy in the spectrum found in Figure 7.15 is

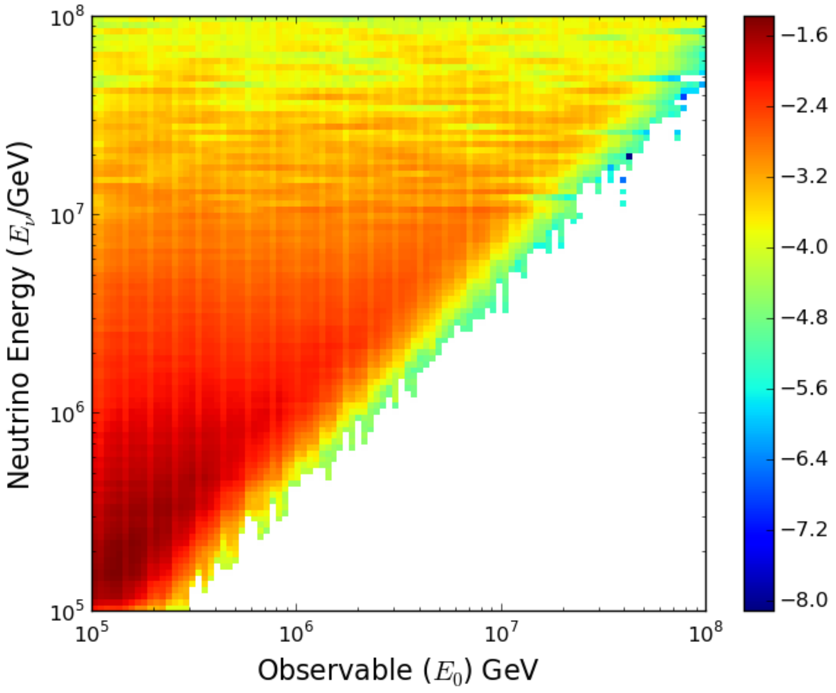


Figure 7.14: The spectrum of incident neutrino energy given the Edepillim reconstructed muon energy. The spectrum is weighted using the astrophysical weights found in [21].Figure 7.15 shows the neutrino energy slice taken for the Edepillim reconstructed energy at $E_0 = 6.3 \times 10^6$ GeV.

$1.0^{+1.1}_{-0.5} \times 10^7$ GeV (10 Pev) whereas the median energy found in [21] using Figure 7.2 is 8.7×10^6 GeV. While there is some variation in the answer due to the different energy reconstruction methods energy reconstruction, both neutrino energies are consistent with one another as there is a large spread in the final probability distribution of these neutrino energies.

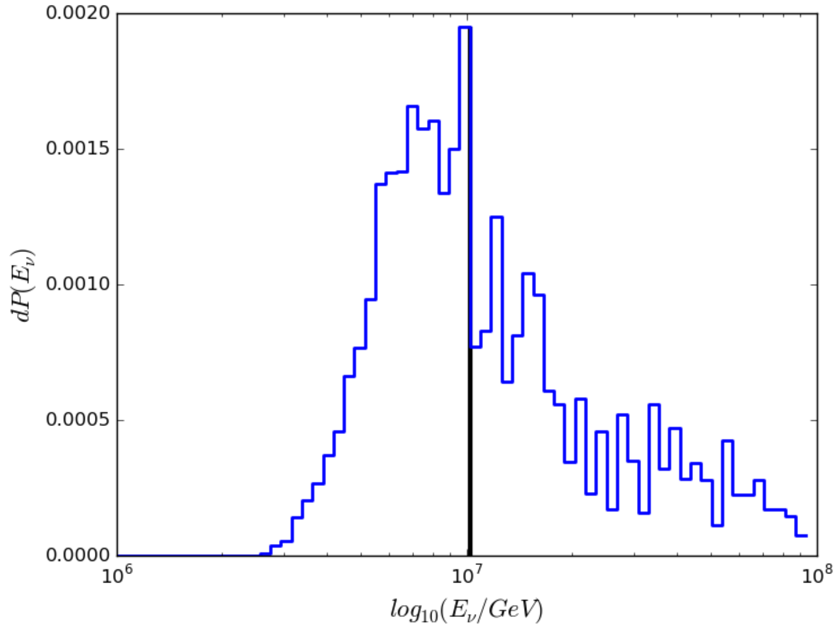


Figure 7.15: The spectrum of the most probable incident neutrino energy, given the Edepillim reconstructed muon energy, for the multi-PeV event that has an Edepillim reconstructed energy of $E_0 = 6.3 \times 10^6$ GeV. The spikes are due to the statistics of the Edepillim reconstruction distribution (Figure 7.13) and the neutrino energy spectrum (Figure 7.12).

7.4 Summary

The highest energy event seen by IceCube to data is a multi-PeV event with a highly stochastic energy loss pattern. Here we have shown that Edepillim has a high resolution for such high energy through-going events by analysing the resolution of a simulation of similar events. The results show a reconstruction resolution better than TE and approaching the reconstruction resolution using Edepillim on true energy losses.

In this chapter we have reconstructed the multi-PeV event's muon energy using Edepillim, based on the optimal resolution setting found using simulation events, and have found a muon energy of $E_0 = 6.3 \times 10^6$ GeV with a median neutrino energy of 1×10^7 GeV assuming a best fit neutrino flux spectrum found in the multi-year analysis in [21]. This is higher than the result published in [21] of 8.7×10^6 GeV but still consistent due to the large uncertainties in neutrino energy from both methods.

Chapter 8

Diffuse Neutrino Energy Spectrum Analysis

The ultimate measure of performance of a muon energy reconstruction is not necessarily the reconstruction resolution in Chapter 4 but the improvement of performance the reconstructed energy method has on a particular analysis. In this chapter we consider the standard upward-going neutrino-induced muon diffuse analysis such as in [9], [30] and [21]. We perform a simplified diffuse analysis on simulations using both the standard truncated energy (TE) and the Edepillim methods, to assess the performance in the overall analysis using each method.

8.1 Diffuse Energy Spectrum

For this analysis a subset of simulations generated for the six year diffuse analysis is used, as outlined in Appendix A. To produce the predictions for the astrophysical and conventional flux, each simulated event in the sample has its muon energy weighted by

$$\text{Weight} = \text{OneWeight} \times E_\nu^{-\gamma} \tag{8.1}$$

where OneWeight is a weighting factor produced during the simulation production which accounts for the energy distribution the events were simulated at [50] and the $E_\nu^{-\gamma}$ is the desired spectrum of neutrino energies. The astrophysical neutrino spectrum assumed in this analysis is $E_\nu^{-2.13}$, which is the best fit flux for the six year diffuse analysis [21]. For the conventional atmospheric flux a single power law of $E^{-3.7}$ is used, which, unlike the spectrum used in the six year analysis, has no angle dependence, but this is a simplification that is acceptable for our purposes. The simulated true muon energy spectrum at the detector can be convolved with the reconstruction distributions, of observable reconstructed energy vs the true muon energy, for different energy reconstructions to create the various energy spectra for the reconstructed energy. In Figure 8.1 the Edepillim reconstruction distribution from Section 6.5.2 for segment length $X = 100\,\mathrm{m}$ has been convolved with the true energy spectrum, that is, each true energy bin has been smeared out by the vertical slice in reconstructed energy from the energy reconstruction distribution. The convolution does not just cause smearing but also creates a bias at energies where the reconstruction distribution is nearly flat along true energy, this is seen in Figure 8.1 as the bump between energies $10^3 - 10^4$ GeV as the true energies below this have been offset to this range of energies due to the shape of the Edepillim reconstruction distribution (see Figure 6.35 as an example which would cause an offset in the convolution). To reduce the effects of outliers, each slice in reconstructed energy has been fitted to a Gaussian. This smearing of true energies into reconstructed energies is done for astrophysical and conventional fluxes to make predictions for the counts of astrophysical and conventional muons for any type of reconstruction.

It is important to note that for the true muon energy distributions there is a large overlap between the astrophysical and conventional distributions with both also having large widths. This means the diffuse fit with different reconstruction methods will not have a substantial improvement in the confidence interval of the fit, due to the fact that the scale of the energy resolution is much smaller than the scale of the widths of the distributions.

This convolution with the true muon energy is done for each reconstructed energy distribution; TE, Edepillim on the true losses, and Edepillim on the Millipede losses. The fluxes for all the energy reconstructions are normalised to the expected number of events of $N_{conv} = 350000$ and $N_{astro} = 500$, which are based on the best-fit results from the six-year diffuse analysis using IceCube data [21]. Due to the convolution with energy reconstruction, some events will be moved outside the energy range used for the likelihood fit, thus the total number of events will be slightly different for each test, as shown in Table 8.1, but this will not affect the fit as the total number of events used in Section 8.2 will be taken as the number of events in the fitted energy range.

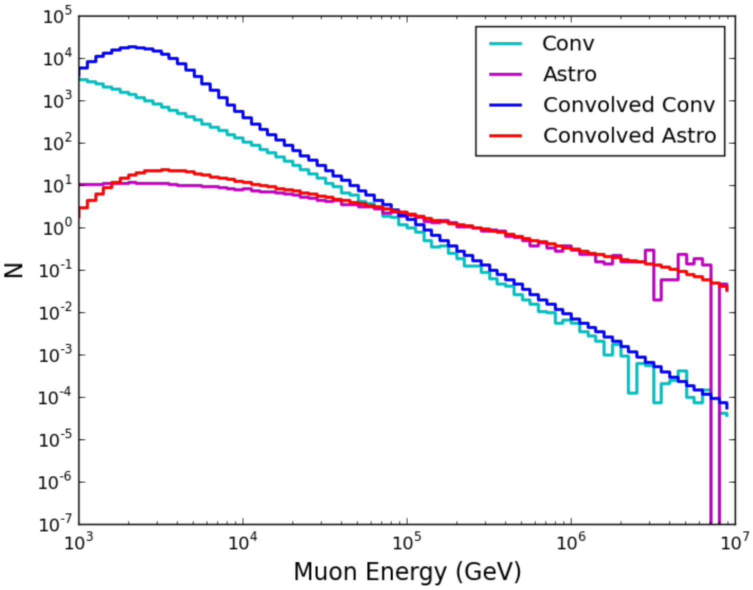


Figure 8.1: The diffuse energy spectrum for the true muon energy (*pink* and *cyan*) and for the Edepillim reconstructed energy (*blue* and *red*). Each Edepillim reconstructed energy distribution was made by convolving the energies with a Gaussian approximation to the reconstructed distribution from Chapter 6. The Edepillim reconstruction causes a bump in the spectrum between muon energies 10^3 to 10^4 GeV which is shown in the *blue* and *red* distribution lines, any true energy energy below 10^3 GeV has been offset to between 10^3 to 10^4 GeV.

8.2 Poisson Likelihood

We perform a simplified diffuse analysis similar to what was outlined in Section 2.5 to test the accuracy of an analysis using the Edepillim muon energy reconstruction against one using the standard TE. We perform a simple one-parameter fit for the strength F, which is the fraction of normalisation of the best-fit astrophysical flux i.e. the total number of astrophysical events will be $N_{astro} = 500 \times F$, and compare each energy reconstruction method by comparing the median one-sigma confidence interval of the parameter F. A Poisson statistic likelihood can be used as follows

$$L = \prod_{i}^{N} \frac{\mu_i^{x_i}}{x_i!} \exp -\mu_i \tag{8.2}$$

where μ_i is the expectation for the count in a muon energy bin i and x_i is the data value for the same energy bin i and N is the number of bins. Taking the natural logarithm yields

$$\ln L = \sum_{i}^{N} x_i \ln \mu_i - \mu_i \tag{8.3}$$

The expectation μ_i in this case is a mixture of the astrophysical and the conventional flux with the astrophysical flux being normalised to a fraction F (the parameter fitted here) of the best-fit flux found in the six year analysis. The total number of events must be preserved, so the expectation in any bin is given by

$$\mu_i = \frac{(N_{total} - N_{astro} \times F)}{N_{conv}} \times \mu_{iconv} + F \times \mu_{iastro} \tag{8.4}$$

where μ_{iconv} and μ_{iastro} are the expectations for conventional and astrophysical given by the flux normalised to the expected number of events for an energy bin i. N_{total} is the total number of muon events, N_{astro} and N_{conv} are the number of astrophysical and conventional muon events for the given flux of the best fit six year analysis in the fitted energy range. As the value of F changes the number of expected astrophysical events will change, which will cause the number of conventional events to change such that the total number of muon events remains the same. The likelihood of each F evaluates how well this mixture fits the observed data. The likelihood in Eq 8.3 can scanned over for different values of F to find a minimum, giving the best fit normalisation of the two fluxes. Wilks' theorem states that the test statistic or likelihood ratio test $-2\log(L/L_0)$ asymptotically will be χ^2 distributed, thus the one-sigma confidence interval on the error of the fit for the normalisation fraction F for each likelihood space will be the values of F where $\log L < \log L_0 + 0.5$, where L_0 is the minimum of the likelihood space.

8.3 Diffuse Analysis on Asimov Data

For a simple analysis to test the relative effectiveness of the two muon reconstruction methods, a full analysis does not need to be done, but the one-sigma confidence interval on the fit for F can be found for each method. Traditionally this is done by generating pseudo-experiments from the simulations and performing a likelihood fit with the representative confidence interval on F being given by the median confidence interval from all the pseudo-experiments.

Instead of generated pseudo experiments, we will use an Asimov data set, which is the Monte Carlo prediction for the assumed true parameter values which, as described in [51], behaves as the median of all the pseudo-experiments. Using this dataset, the likelihood function will minimise to the original input model parameters and the median one-sigma confidence interval on the model parameter can be found. The Asimov dataset for this analysis is the combined astrophysical and conventional energy spectrum for each reconstructed muon energy from the six year analysis.

Using the Asimov datasets the Poisson likelihood from Eq 8.3 was found for each type of energy reconstruction with the likelihood as a function of varying F shown in Figure 8.2, with the one-sigma confidence intervals shown. The confidence interval width for the reconstruction method of Edepillim using the true energy losses is only slightly wider than that using the true muon energy. This is expected as Edepillim on true losses has been shown to have a good ALF reconstruction resolution with values at approximately $\sigma_{ALF} = 0.15$ for all target true energies. The interval for Edepillim reconstruction on Millipede losses is only slightly smaller than the TE interval and they are very hard to distinguish in Figure 8.2. All of these results are summarised in Table 8.1.

The connection between these diffuse results and resolution results in previous chapters can be made by comparing the ALF reconstruction resolution (for some appropriate representative energy) to the width of the likelihood at the one-sigma level. The ALF value at muon energy $E_{TRUE} = 10^5$ GeV was chosen, as it is the energy where the crossover between the astrophysical and conventional energy distribution happens and as such the resolution at this energy is expected to have a significant impact on the fit. The ALF and confidence interval widths are shown for different reconstruction types by the points

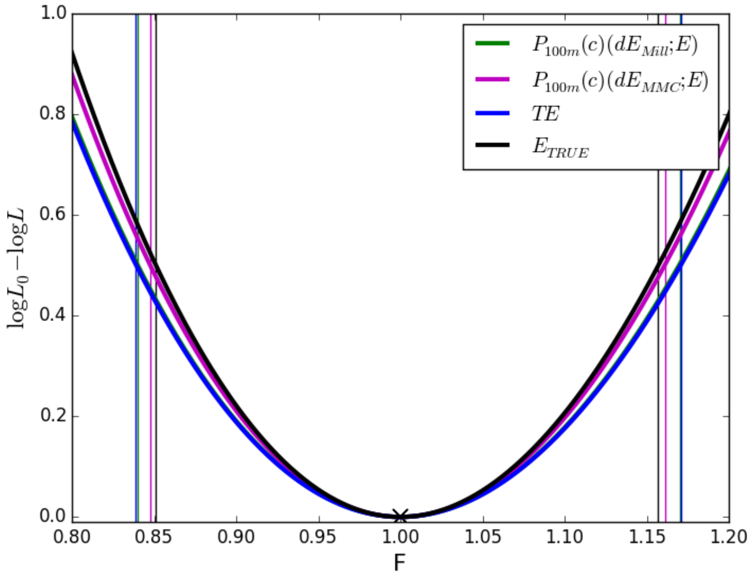


Figure 8.2: The log likelihood fit to Asimov data for TE (*blue*) and Edepillim on Millipede reconstructed losses (*green*) and Edepillim on the true energy losses (*pink*). The likelihood values have been adjusted for display on the same plot by adjusting all values by the minimum likelihood value to place the minimum at zero. The one-sigma confidence interval is shown for each plot, with the width giving the error on F for each energy reconstruction method. The width for the Edepillim reconstruction on true losses is the smallest for all the energy reconstructions, with its likelihood approaching the result of using the true losses. The likelihood for Edepillim reconstruction on Millipede reconstructed losses is indistinguishable from the likelihood for TE, showing there is little difference between the two reconstructions for this simplified diffuse fit.

in Figure 8.3. Also shown are the results if the true muon energy distributions are used, as this would give the best possible one-sigma interval result.

The black line shows how the one-sigma interval varies with the ALF resolution for an idealised reconstruction distribution, in this case a Gaussian with width $\sigma_{orig} = 0.1$ for each true energy reconstruction, centred at the true muon energy, with σ_{orig} being scaled between 0 and 0.3. The Gaussian reconstruction distribution is then convolved with true energy, and for each scaling of the Gaussian width in reconstructed energy the ALF and diffuse fit confidence interval width is obtained to create the black line in Figure 8.3. This shows how the confidence interval vs ALF resolution varies for a series of idealised reconstruction distributions. The plot also shows lines which represent the ALF and diffuse fit results for taking the reconstruction distribution fitted Gaussians for each energy and scaling these by a factor in the range 0 to 3. For each reconstruction this results in the coloured lines (*pink*, *green* and *blue*) in Figure 8.3. This retains the general form of the spread of true energy into reconstructed energy as a function of true energy, but then improves or worsens the resolution by a scaling factor.

The connection between the one-sigma confidence interval and the ALF resolution in Figure 8.3 is clear with a rise in the ALF reconstruction resolution resulting in an increase in the confidence interval width for the diffuse likelihood. This is not a linear increase, for example a change of ALF resolution in intervals of 0.05 for σ_{ALF} from 0.15 to 0.3 results will result in confidence interval increases of 1.2%, 1.9% and 2.4% , which shows a parabolic relationship. The limiting case of fitting using the true muon energy is also shown (*black dot*), at this point the ALF resolution is zero and all true muon energies are exactly known. The result for the Edepillim energy reconstruction on true losses (*pink*) is shown with its own scaling line. The scaling line for the fit to Edepillim reconstruction on true losses mimics that of the pure Gaussian line. This indicates that the Edepillim reconstruction on true losses is equivalent to a Gaussian of a similar constant sigma in reconstructed vs true energy. The Edepillim energy reconstruction on Millipede losses (*green*) and TE (*blue*) also have lines representing the scaling of fitted sigmas, as the sigma is reduced these lines approach the pure Gaussian line and all lines meet at the black dot when scaled to zero sigma.

Of interest is that here the Edepillim reconstruction on Millipede losses performs better than TE for both the ALF resolution and the likelihood width. The reason for the improvement on the ALF reconstruction relative to Chapter 5 is that the fitting of a Gaussian to the reconstruction distribution has reduced the impact of outliers in the resolution, which has a larger effect on the Edepillim resolution. This improvement also translates into the diffuse likelihood with Edepillim having a smaller confidence interval width by $\Delta width = 0.002$, but this is not a significant effect.

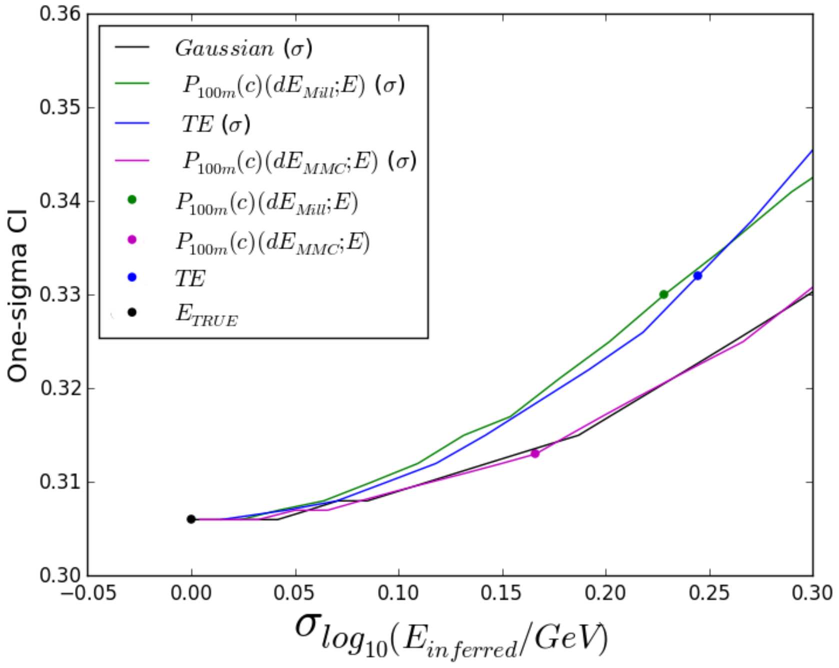


Figure 8.3: The width of the likelihood at one-sigma vs the ALF reconstructed resolution at $E_{TRUE} = 10^5$ GeV for the reconstructed muon energies Edepillim on Millipede (*green dot*), Edepillim on true losses (*pink dot*), TE (*blue dot*) and the true muon energy (*black dot*). Shown for comparison are the results of the same analysis on true muon energies convolved with a Gaussian with a scaling σ between 0 and 0.3 (*black*). As the value of σ increases the resolution and width both increase in relation to each other. The lines (*green*, *pink* and *blue*) show the results when the σ of the Gaussian fit for each energy reconstruction is scaled by a factor between 0 and 3. Each scaling line shows the trend for each reconstruction, when at the smallest ALF resolution each reconstruction approaches the same confidence interval width.

The expectation of Edepillim performing better with a selection on long path lengths based on the results in Chapter 6 can be tested by creating a subset of events which only has events with path lengths greater than 1000 m. The energy spectrum of different energy reconstruction methods are made using a subset of the events with long path lengths and by convolving with a reconstruction resolution obtained for the same path length selection. The convolved energy distribution plot for long events is shown in Figure 8.4 for the Edepillim muon reconstruction on Millipede losses and for TE. Note the energy spectrum for both reconstructions is similar at high muon energies, and this will result in little difference in the diffuse fit. The likelihood scan in Figure 8.5 shows the diffuse fit for the different energy reconstructions as well as the baseline of using the true muon energy, E_{true} (*black*). The width of the one-sigma confidence interval for the Edepillim reconstruction on true energy losses (*pink*) is smallest and closest to the true muon energy. In this case the Edepillim reconstruction on Millipede energy losses (*green*) is noticeably smaller than that of TE (*blue*). The likelihood for Edepillim on Millipede losses could approach that of Edepillim reconstruction on true losses with improvements to the Millipede reconstructed energy losses.

The relation between the one-sigma confidence interval width in the diffuse likelihood and the ALF reconstruction resolution for the various reconstructions, for long events, is shown in Figure 8.6. As for Figure 8.3 the results of convolutions with a Gaussian with changing sigma are also shown, the different energy reconstructions also follow the same pattern as in Figure 8.3, but all the results for the diffuse fit have increased the one-sigma interval fit due to fewer events being in the energy spectrum. This is effectively like running the detector for a shorter time, which results in a less well constrained astrophysical flux. Using only long events has caused the difference in the likelihood width between Edepillim on Millipede losses and TE to be slightly larger with $\Delta width = 0.004$. This means that there is very little difference in the performance of the Edepillim reconstruction on Millipede losses and TE, as even in the case of using the best reconstruction resolution for both there is only a difference of the order of 10^{-3} in the confidence interval width (see Table 8.1) for the diffuse analysis. It is also clear in the long events case that the Edepillim reconstruction on true energy losses resembles the Gaussian baseline, as both lines for the variations in the Gaussian sigma are similar.

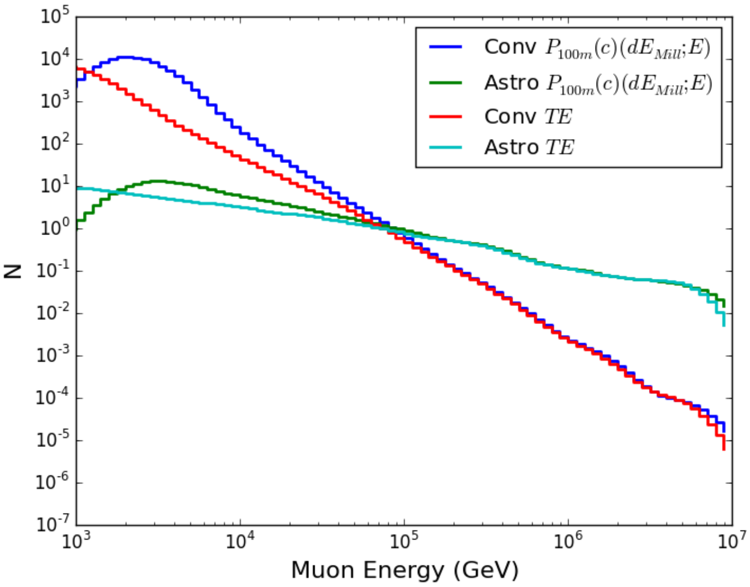


Figure 8.4: The diffuse energy spectrum made from the convolution of the true muon energy with the Edepillim reconstructed energy distribution (*blue* and *green*) and the TE distribution (*red* and *cyan*) for long events with path length $L > 1000\,\text{m}$. The Edepillim and TE reconstructed energy distributions was made by convolving the energies with a Gaussian approximation to the reconstructed distribution from Chapter 6. The spectra for both Edepillim and TE are very similar at high energies, including near the intersection energy where the astrophysical flux becomes more dominant.

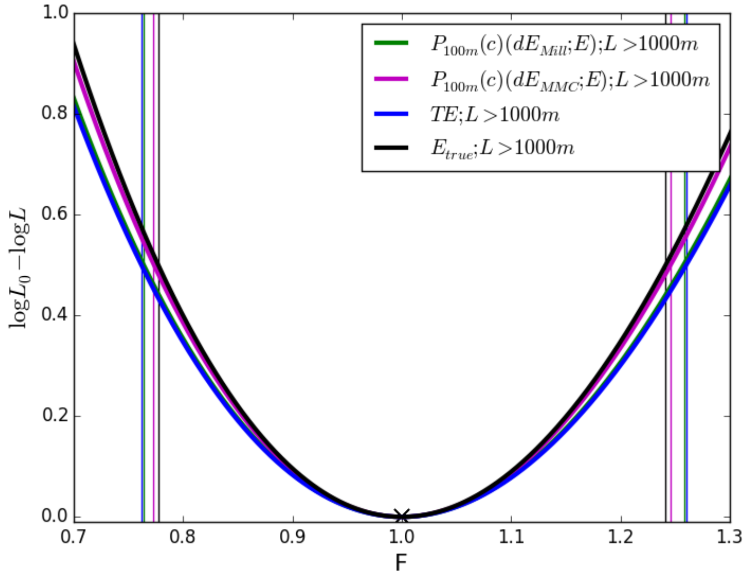


Figure 8.5: The log likelihood fit to Asimov data for TE (*blue*) and Edepillim on Millipede reconstructed losses (*green*) and Edepillim on the true energy losses (*pink*) where only events with path lengths $L > 1000\,\mathrm{m}$ are included in the flux. The likelihood values have been adjusted for display on the same plot by adjusting all values by the minimum likelihood value to place the minimum at zero. The one-sigma confidence interval is shown for each plot, with the width giving the error on F for each energy reconstruction method. The TE likelihood lies on top of the Edepillim on Millipede losses likelihood thus there is no clear difference in using TE or Edepillim with long path lengths.

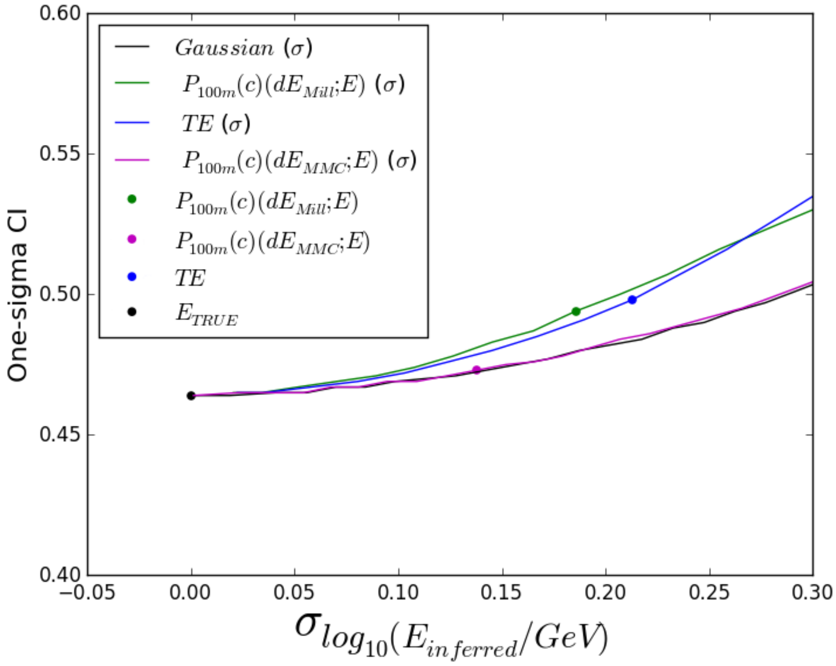


Figure 8.6: The one-sigma interval width of the likelihood vs the ALF reconstructed resolution at $E_{TRUE} = 10^5$ GeV for the reconstructed muon energies Edepillim on Millipede (*green dot*), Edepillim on true losses (*pink dot*), TE (*blue dot*) and the true muon energy (*black dot*). Shown for comparison are the results of the same analysis on true muon energies convolved with a Gaussian with a scaling σ between 0 and 0.3 (*black*). As the value of σ increases the resolution and width both increase in relation to each other. The lines (*green*, *pink* and *blue*) show the results when the σ of the Gaussian fit for each energy reconstruction is scaled by a factor between 0 and 3. Each scaling line shows the trend for each reconstruction, when at the smallest ALF resolution each reconstruction approaches the same confidence interval width.

Energy Reconstruction	Events	One-sigma CI	σ_{ALF}
TE	All(198503)	0.332	0.244
$P_{100m}(c)(dE_{Mill}; E)$	All(198691)	0.330	0.228
$P_{100m}(c)(dE_{MMC}; E)$	All(198530)	0.313	0.166
E_{TRUE}	All(198691)	0.304	0.000
$TE; L > 1000m$	Long(114794)	0.498	0.213
$P_{100m}(c)(dE_{Mill}; E); L > 1000m$	Long(114954)	0.494	0.186
$P_{100m}(c)(dE_{MMC}; E); L > 1000m$	Long(114952)	0.473	0.138
$E_{TRUE}; L > 1000m$	Long(114954)	0.461	0.000
$TE; L < 1000m$ $+P_{100m}(c)(dE_{Mill}; E); L > 1000m$	All (83705+114954)	0.330	$\sigma_{short} = 0.263$ $\sigma_{long} = 0.186$

Table 8.1: Summary of the results for the different muon energy reconstructions for the diffuse likelihood fit and the ALF reconstruction resolution. Shown are the values for fits using all events, a subset of long path length events and a combined diffuse fit with TE used for short events and Edepillim on Millipede losses for long events. The difference in the one-sigma confidence interval for TE and Edepillim on Millipede losses is only of the order of 10^{-3} and so is not a significant difference, this also applies to fits done using all events and long events. There is also little difference when using the combined likelihood fit.

8.4 Combined Reconstruction

A final study was done using a combined likelihood for TE and Edepillim on Millipede losses where TE is used for events with path length below 1000 m (short events) and Edepillim's reconstructed energy using Millipede losses is used for events with path length greater than 1000 m (long events). The two reconstructions must still be treated as separate reconstructed energy distributions as the reconstructed energies are not interchangeable and must be normalised to the expected number of events for each sub-sample. For the case of using TE on events with path length $L < 1000\,\mathrm{m}$ in the tested energy range there are, after weighting, 223 astrophysical events and 83482 conventional events. For long events with path length $L > 1000\,\mathrm{m}$, the Edepillim resolution on Millipede losses is used with 239 astrophysical events and 114715 conventional events. There is now half the number of astrophysical events in the short TE distribution and the other half in the astrophysical Edepillim energy distribution. As in the previous section, the number of events are not the exact numbers the true energy losses were normalised at, this is due to the convolution moving some events to outside the fitted energy range. The combined likelihood requires a likelihood for each energy reconstruction's spectrum to be calculated separately then combined via

$$\begin{aligned}\ln L = &\sum_i^N (x_{i\mathrm{TE}} \ln \mu_{i\mathrm{TE}} - \mu_{i\mathrm{TE}}) \\ &+ \sum_i^N (x_{i\mathrm{ED}} \ln \mu_{i\mathrm{ED}} - \mu_{i\mathrm{ED}})\end{aligned} \tag{8.5}$$

where the data points are from the short event's TE distribution $x_{i\mathrm{TE}}$ and the long events from the Edepillim energy distribution $x_{i\mathrm{ED}}$. The expectation equation is based on that in Eq 8.4, but with N_{total} now the number of events expected for each spectrum, and N_{astro} and N_{conv} for each distribution as follows

$$\mu_{i\mathrm{ED}} = \frac{(N^{ED}_{total} - N^{ED}_{astro} \times F)}{N^{ED}_{conv}} \times \mu_{i\mathrm{ED}conv} + F\mu_{i\mathrm{ED}astro} \tag{8.6}$$

Here $N^{ED}_{total} = 114954$, $N^{ED}_{astro} = 239$ and $N^{ED}_{conv} = 114715$ to maintain the consistent total for the Edepillim reconstruction energy distribution. For the TE distribution the expectation is given by

$$\mu_{i\mathrm{TE}} = \frac{(N^{TE}_{total} - N^{TE}_{astro} \times F)}{N^{TE}_{conv}} \times \mu_{i\mathrm{TE}conv} + F\mu_{i\mathrm{TE}astro} \tag{8.7}$$

Here the number of events correspond to $N^{TE}_{total} = 83705$, $N^{TE}_{astro} = 223$ and $N^{TE}_{conv} = 83482$ for the fitted energy range. While the common best fit F is being found simultaneously across Eq 8.5, Eq 8.6 and 8.7, the two separate normalisation conditions are being enforced.

The results of the combined energy reconstruction in Figure 8.7 show a very slight improvement in the width of the confidence interval for the combined events over using TE or Edepillim on Millipede losses on all events. As with the case of the long events in Figure 8.5, the improvement in the confidence interval with the Edepillim method is very small, with the change in the confidence interval widths occurring on the order of 10^{-3} (see Table 8.1) which could not be considered to be a clear indicator of improvement in the likelihood fit. The reason for the lack of improvement with combined likelihood is due to the results shown in the previous section where for the subset of long events there was little change in the results for TE and Edepillim on Millipede losses. Because the long event energy distributions are so similar there is no clear improvement in the diffuse fit.

8.5 Summary

In this chapter we have tested the Edepillim method against TE using a simulated diffuse analysis. The analysis used a Poisson likelihood for the expectation of the flux given the data, and to measure sensitivity of each method we used an Asimov data set to represent a median of the pseudo experiments and used the one-sigma confidence interval on the astrophysical flux as a comparison measure.

The results, as summarised in Table 8.1, showed that there is little difference between the diffuse flux fit on TE or Edepillim reconstruction on Millipede losses on all events, with Edepillim performing only slightly better. The results of the diffuse fit on Edepillim reconstructed events on true losses show a smaller confidence interval, again by only a small value, but does show the potential for Edepillim to improve on the already existing energy resolution in the future with improvements on reconstructed energy losses from Millipede.

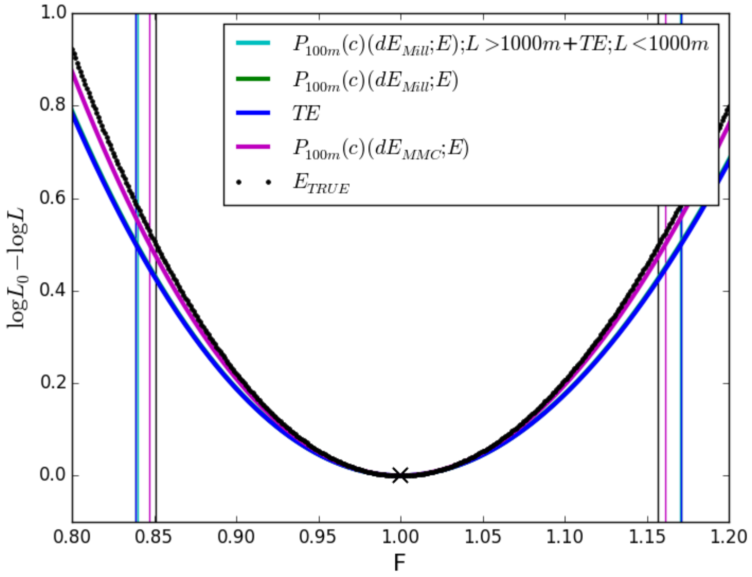


Figure 8.7: The log likelihood fit to Asimov data for the combined energy reconstruction (*cyan*) of TE on events with path length $L < 1000\,\mathrm{m}$ and Edepillim reconstructed energy using Millipede energy losses on events $L > 1000\,\mathrm{m}$, with TE (*blue*) and Edepillim on the true energy losses (*pink*) for all events shown for comparison. The likelihood values have been adjusted for display on the same plot by adjusting all values by the minimum likelihood value to place the minimum at zero. The one-sigma confidence interval is shown for each plot, with the width giving the error on F for each energy reconstruction method. The combined likelihood has only a very slightly smaller confidence interval width compared with using TE or Edepillim using Millipede losses on all events.

The connection between the ALF reconstructed resolution and the confidence interval width is demonstrated as a parabolic relation with improvements observed in ALF resolution resulting in much smaller improvements in the error on the diffuse fit. For example, reducing the spread in the reconstruction resolution from 0.3 to 0.25 leads to only a 2.4% improvement in the one-sigma confidence interval width, however a further reduction from 0.25 to 0.2 only leads to only a 1.9% improvement in the one-sigma confidence interval width. This indicates that improvement in the energy reconstruction resolution will only result in smaller improvements in the accuracy of the diffuse fit.

The small improvement in the diffuse fit for different reconstructions can be explained by observing the large widths of the energy spectra for astrophysical and conventional neutrinos. There is a high overlap in both distributions so any improvement in energy resolution will not have a substantial impact on the overall fit, as the energy resolution is much smaller than the width.

The work in this chapter indicates only small improvements are to be expected by using Edepillim reconstructed energy losses in a diffuse fit analysis. However this is not the only use for an Edepillim muon energy reconstruction. For example in Chapter 7 we described its use to reconstruct a single high energy muon event, and its use may improve other types of analyses.

Chapter 9

Conclusion

The IceCube neutrino observatory detects neutrinos from astrophysical and atmospheric sources. Each has a unique energy spectrum and the separation between these spectra can be done with a diffuse analysis. The detection of neutrinos relies on the detection of secondary particles, muons. The energy of the muons needs to be reconstructed with a high reconstruction resolution in order to be used in analyses.

As described in Chapter 5, we have developed a muon energy reconstruction method called Edepillim, which uses a muon's energy loss pattern to perform a maximum likelihood fit over the probabilities of a muon's energy loss for a given muon energy. This method was designed and here optimised to be used in the neutrino detector IceCube. In Chapter 4 we have outlined the method for measuring the reconstruction resolution, which is then used in all later chapters. In an idealised case (Chapter 5) the Edepillim method reconstructs the muon energy with a very good reconstruction resolution when using true energy losses. The reconstruction resolution improves with the amount of information used in the likelihood fit, with improvements seen when using smaller segments lengths (i.e. having high resolution energy losses) and longer path lengths.

We have also tested the Edepillim method on an IceCube simulation of muon events in Chapter 6 and found that the reconstruction resolution for Edepillim is much worse when the reconstructed energy losses are used, compared with using the true energy losses. The reconstruction of the muon energy loss pattern is performed by Millipede, but studies of the Millipede energy loss patterns show that the reconstructed energy losses differ from the true energy losses in ways that strongly affect Edepillim. Millipede reconstructed patterns have an abundance of zero energy losses which cause problems with the Edepillim reconstruction. When the track has over 60% of its losses as zero losses the Edepillim reconstruction will return the sum of the energy losses as the reconstructed energy, which is due to the likelihood becoming stuck at that energy, see Section 6.5. Additionally, Millipede also has the issue of having inaccurate energy losses, the most obvious being losses that are orders of magnitude larger than any true energy loss, that are reconstructed outside the detector volume. Many of these losses can be removed by only using the energy losses that occur within the detector volume. The accuracy of the energy losses can be improved by combining the energy loss bins to create large segment lengths, which results in a large improvement of the Edepillim reconstruction resolution.

The Edepillim reconstruction uses a PDF which gives the probability of a muon energy loss given a muon energy, but due to the construction of the PDF it does not contain a zero energy loss bin. In order to account for the zero energy losses that occur in the energy loss pattern, the continuous loss per segment length can be included. This allows for a PDF to be constructed that contains both continuous and stochastic energy losses, and thus every occurrence of zero stochastic loss will still correspond to a probability on the PDF.

At the end of Chapter 6 two selection criteria for muon events were discussed to improve the Edepillim resolution. The best selection cut was found to be on the path length of the muon event, with the Edepillim reconstruction on events with a path length greater than 1000 m having a better reconstruction resolution than TE.

In Chapter 7 the Edepillim reconstruction method was performed on the highest deposited energy event detected by IceCube to date. The Edepillim reconstruction was based on the method that resulted in the best reconstruction resolution for a simulation of similar events, which is a simplified version of the method discussed in Chapter 6. Using this simulation resulted in a better reconstruction resolution than in Chapter 6, this is due to the events in this simulation of the specific event being high energy events with well reconstructed energy losses. The Edepillim muon reconstructed energy for the event was $E_0 = 6.3^{+1.9}_{-1.1} \times 10^6\,\text{GeV}$ which was higher than that of the TE ($E_0 = 4.5 \pm 1.5 \times 10^6\,\text{GeV}$) but within the margins of error. The neutrino energy ($E_\nu = 10\,\text{PeV}$) for the event using the Edepillim reconstruction distribution and the best fit neutrino spectrum from [21] is also higher than the published results in [21].

Finally in Chapter 8 a diffuse likelihood fit (a simplified version of the published likelihood six year diffuse fit [21]) was done using the different reconstruction types (TE, Edepillim reconstruction on true losses and Edepillim reconstruction on Millipede losses) to test the performance of Edepillim on a diffuse analysis. The reconstructed energy distributions were convolved with a true muon energy spectrum to obtain the astrophysical and conventional energy distributions expected for each method. To smooth the spectrum the reconstructed distributions for each true energy were fitted to a Gaussian. A likelihood fit was performed on an Asimov dataset which was used to find the one-sigma confidence interval on the fit of the fractional normalisation of the astrophysical flux. The width of this interval was compared to the ALF resolution that was used as a measure of resolution in earlier chapters. The two values are connected by a parabolic relationship with equal improvements in the ALF resolution corresponding to progressively smaller changes in the confidence interval width. The results show the best confidence interval is for the Edepillim reconstruction on true losses, with the confidence interval for Edepillim on Millipede losses, giving a slightly better result than TE. The likelihood fit was also tested for a subset of events with long path lengths, with similar relative results for the confidence interval compared with that using all events. Finally a combined likelihood was done using TE on short path length events and Edepillim on Millipede losses for long path length events, with no significant change in results. The Edepillim energy reconstruction result only provides a minor improvement on the result of the fit due to the fact that the true muon energy distributions for astrophysical and conventional already have a very large width, and a large overlap, thus even with better energy resolution there is not much change to the diffuse fit results. This happens as the scale of the energy resolution is much smaller than the scale of the width of the energy spectra.

The results of the diffuse likelihood fit analysis indicate that there are only minor improvements to be made by using the Edepillim reconstruction on Millipede events. With differences in the one-sigma confidence intervals being very small between even the Edepillim reconstruction and that of using the true muon energies, this indicates that there are only minor improvements that can potentially be made by improvement in energy reconstruction. However, this does not exclude Edepillim from yielding improvements in other analyses, or from being used in cases of reconstructing a specific muon event such as in Chapter 7, where Edepillim worked well since the losses for that particular event were found to be especially well reconstructed.

Further improvements in the reconstructed energy losses coming from Millipede will produce improved reconstruction resolution for Edepillim. This could be achieved by the joint likelihood approach of using Edepillim to constrain the Millipede reconstructed energy losses, during the Millipede reconstruction. In the future Edepillim could also be used for other applications such as energy reconstructions of tau events, or starting neutrino events.

Appendix A

Muon Event Simulation Datasets

In this work several different muon event simulations were used. Table A.1 describes these and gives the corresponding chapters each simulation set was used in.

Muon Event Simulation		
Chapter	Energy Range	Description
Chapter 5	$10^2 - 10^{10}$ GeV	A simulation of muon event energy losses that was made specifically for this work to test Edepillim in a true energy loss case. The simulation was done using the open source software version of PROPOSAL, with muons simulated in-ice but with no other detector conditions imposed. There are equal numbers of muon events per log energy bin.
Chapter 6	$10^2 - 10^7$ GeV	An IceCube produced simulation using muons started just outside the detector, with equal numbers of muon events per log energy bin.
Chapter 7	$10^5 - 10^8$ GeV	An IceCube simulation of events with similar trajectory and energy as the highest energy muon event detected to date. There are equal numbers of muon events per log energy bin. Produced using a muon simulator that injected muons just outside the detector.

Muon Event Simulation continued		
Chapter	Energy Range	Description
Chapter 8	$10^2 - 10^8$ GeV	An IceCube simulation of diffuse neutrino events used for the analysis in [21]. Produced using a neutrino simulator with muon production occurring from large distances outside, to within, the detector volume.

Table A.1: Table describing the various simulations used throughout this work.

Appendix B

Truncated Energy Standard for Comparison

It was reported in [34] that the best version of the truncated energy (TE) reconstruction for use on high energy simulations is the AllBINs method. However after a reconstruction resolution analysis was performed on the simulation from Chapter 5 AllDOMs was found to have the best resolution as shown in Figure B.1. Thus for Chapters 6, 7 and 8 all comparisons to TE in was done with the AllDOMs reconstruction method with its reconstruction distribution as shown in Figure B.2.

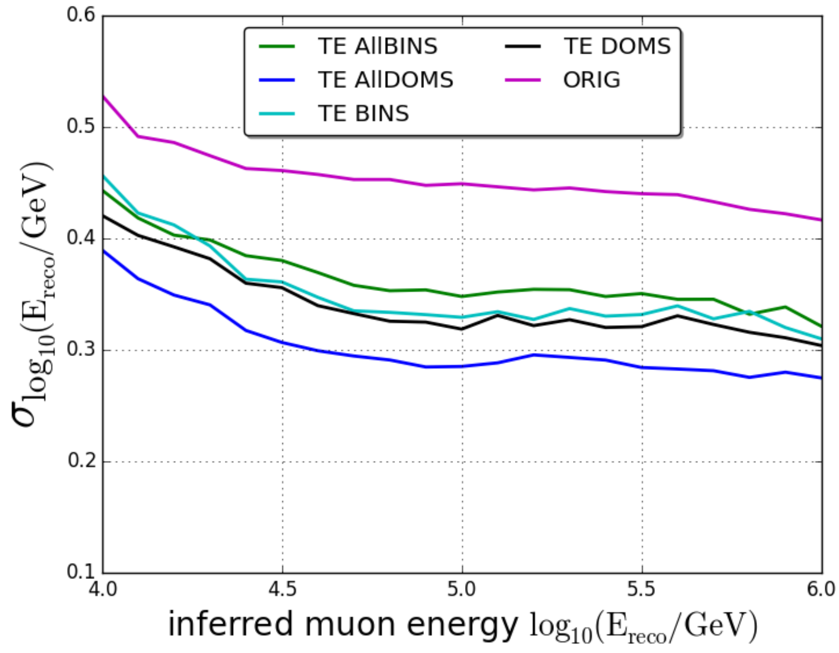


Figure B.1: The ALF reconstruction resolution at target true muon energies, for the different TE versions used on IceCube simulation. TE using AllDOMs has the clear best resolution and so will be the reconstructed energy used for all comparisons. For a simulation of muons used in Chapter 6.

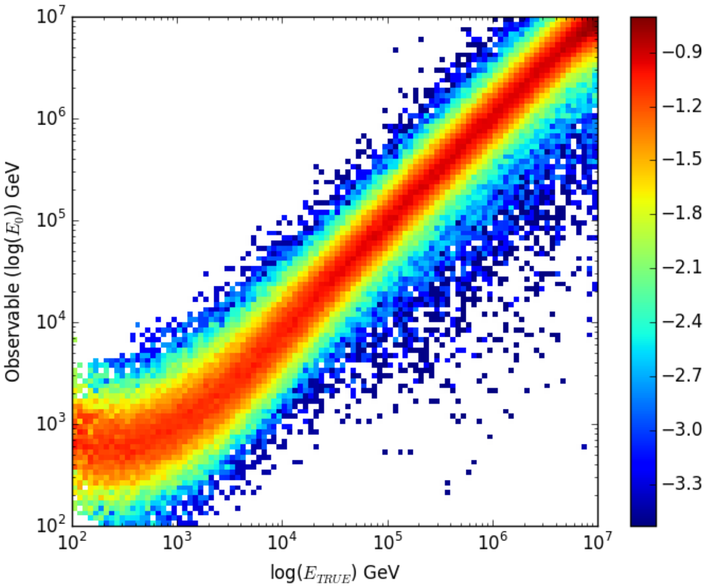


Figure B.2: The reconstruction distribution of observable muon energy E_0 vs muon true energy E_{TRUE} for TE. In this case E_0 is the truncated mean energy loss rate that is calculated using the ALLDOMs method as outlined in Section3.4. For the simulation of muons that was used in Chapter 6.

Appendix C

Smoothing of PDFs using B-Splines

The results of the Edepillim energy reconstruction using the binned PDF described in Chapter 5 can result in a fluctuating likelihood shape which creates a difficulty finding a minimum in the space. When a minimiser is used on the space these fluctuations with affect the performance of the minimiser, as it could interpret these local minima as the global minimum. To avoid this a smooth likelihood space is required, to achieve this it is possible to use an interpolation fitter on the PDFs. As such the probabilities would then be unbinned and produce a smooth likelihood space.

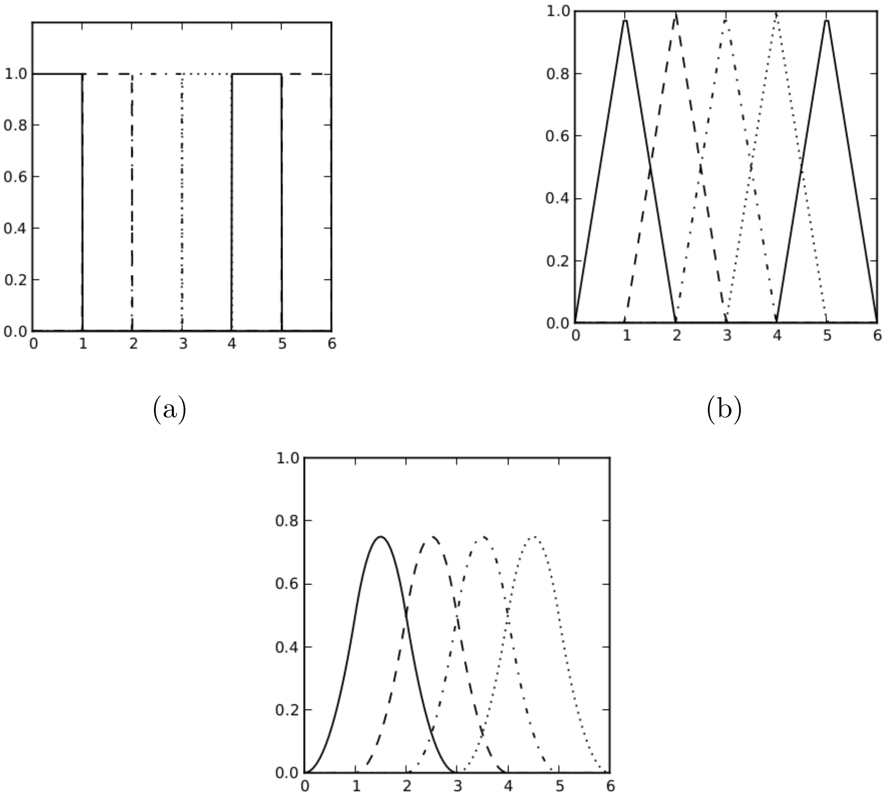
There is a library already in place in the IceCube software which performs this function, `photospline`. The library `photospline` uses a penalised spline technique to process and evaluate the B-spline representation of the data.

A Basis Spline (B-spline) in its simplest form is a unit box shown in Figure C.1a which takes the value of one between the defined knot points $k = 0$ and $k = 1$. This is a basic step function

$$B_0(x) = \begin{cases} 1 : k_0 < x < k_1 \\ 0 : \text{otherwise} \end{cases} \tag{C.1}$$

The order of the spline can be increased such that the B-spline approaches a Gaussian, Figure C.1c, see [52] for more details.

Mathematically a spline is a numeric function that is piecewise-defined by polynomial functions, and which possesses a sufficiently high degree of smoothness at the points where the polynomial pieces connect. These connection points are known as knots and are user defined points that the function fits to, for every spline there will be a knot vector denoted as $\vec{k}$. Between each knot in $\vec{k}$ a polynomial B-spine function is fitted and the final function will be a linear combinations of these B-splines [53].



(a) (b)

(c)

Figure C.1: An example of B-splines with (a) Order-0 (b) Order-1 (c) Order-2 [52]

For higher orders it is possible to evaluate B-splines using a recurrence relation as it requires no special arrangements in cases of multiple knots and does not suffer loss of significance. Such a relation allows the calculation of B-splines that can be combined to form the overall spline function [54].

The selection of knots in the vector $\vec{k}$ is important to ensure the B-splines have enough coverage for adequate linear combination. A one dimensional example of a function fitted to a random set of data is shown in Figure C.2. In Figure C.3 a smaller range of knots is used which affects the fit. Knots should ideally extend beyond the range of the data as the fit will want to tend towards zero at the edge of $\vec{k}$.

In Figure C.3b the number of knots in $\vec{k}$ is reduced, therefore the are fewer B-splines in the fit, this has a noticeable effect between $-10 < x < 5$ as the fit does not follow the data. Thus the number of knots must be large enough to allow enough B-splines to be fitted, but not so large as to significantly slow computational time.

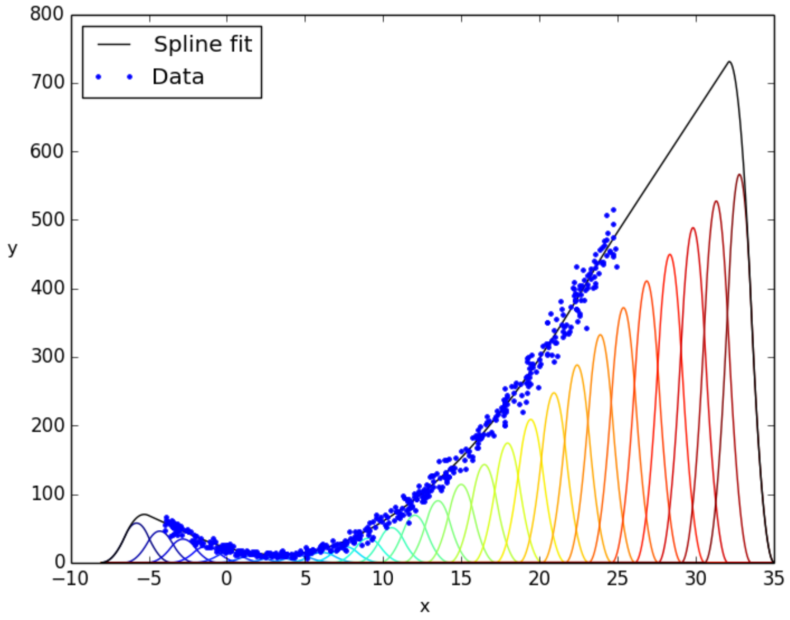
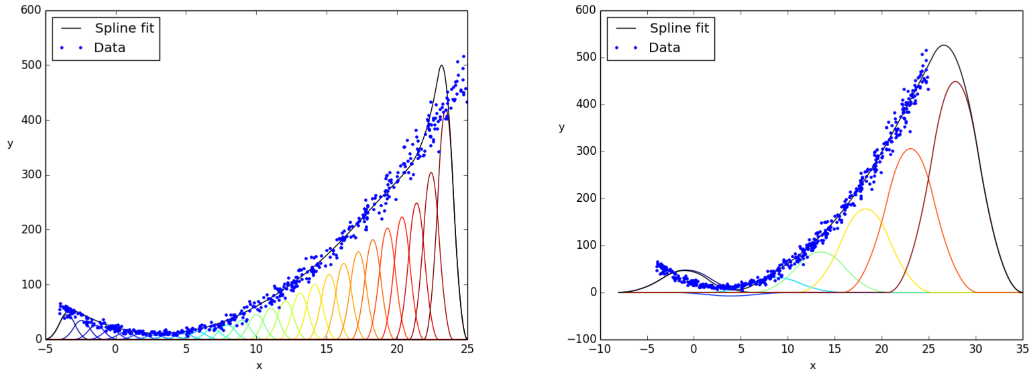


Figure C.2: An example of a set of random data fitted with a series of B-Splines to produce a spline fit.

The B-spline combination fit for an example PDF, given in Figure C.4, is shown in Figure C.5, the knots were defined as a matrix of points along the observable energy and true energy, between $10^0 < E_0 < 10^9$ and $10^{-6} < E_{TRUE} < 10^0$. The knot range extends beyond the required PDF range to ensure there is enough coverage for the fit, to avoid problems as demonstrated in Figure C.3a.

The overall effect of a fitted PDF can be most accurately seen in the final reconstruction resolution that results from it. In Figure C.7 the resolution from using a binned and splined (unbinned) PDF is shown, with both the true energy loss pattern (MMC) and the Millipede reconstructed energy loss pattern (Mill) with an improvement in resolution most seen for the reconstruction on Millipede energy losses.



(a) (b)

Figure C.3: An example of splines with variations on the knots assigned to each fit (a) the knots do not cover adequate range to fit to all of the data, (b) there are not enough knots to perform the fit resulting in a worse fit at the lower end.

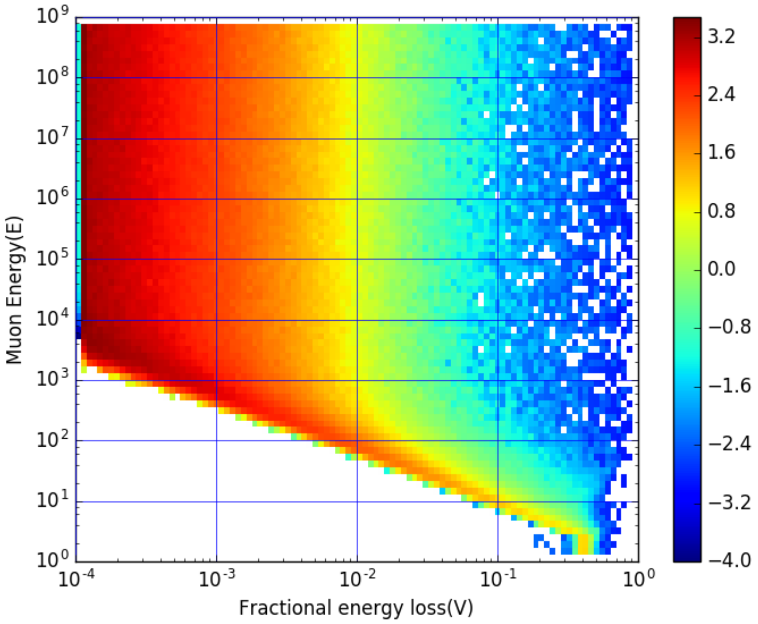


Figure C.4: Probability distribution function built from simulated losses of segment length $X = 10\,\mathrm{m}$ binned along fractional energy loss $V = E/dE$, displayed in the log of probability.

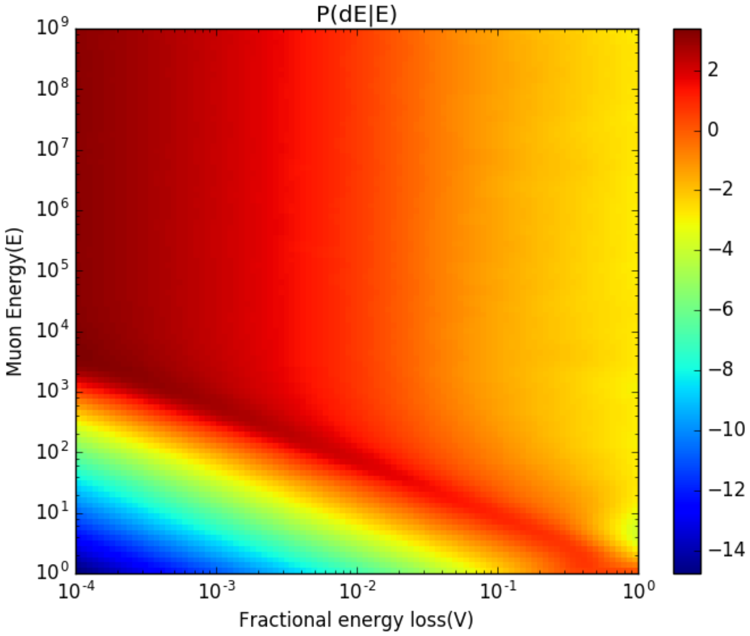


Figure C.5: Probability distribution function built from simulated losses of segment length $X = 10\,\mathrm{m}$ binned along fractional energy loss $V = E/dE$ (Figure C.4), displayed in the log of probability. This PDF has been spline fitted with 200 knots defined between $10^0 < E_0 < 10^9$ and $10^{-6} < E_{TRUE} < 10^0$.

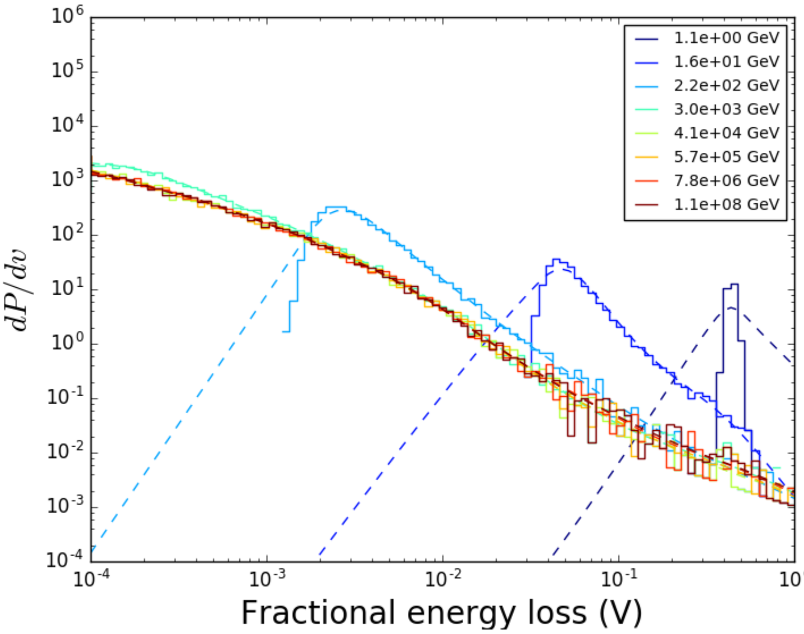


Figure C.6: Particular energy slices of the binned and spline fitted PDF in Figure C.5. The solid lines are the binned PDF, the dashed lines are the spline fitted PDF for the same muon energy. When there are empty probability bins for the PDF the spline fitter fits a PDF with the same trend. The spline fits reproduce the trend of all the muon energy bins with the lower muon energies having fewer bins so there is more extrapolation of the trend for the PDF.

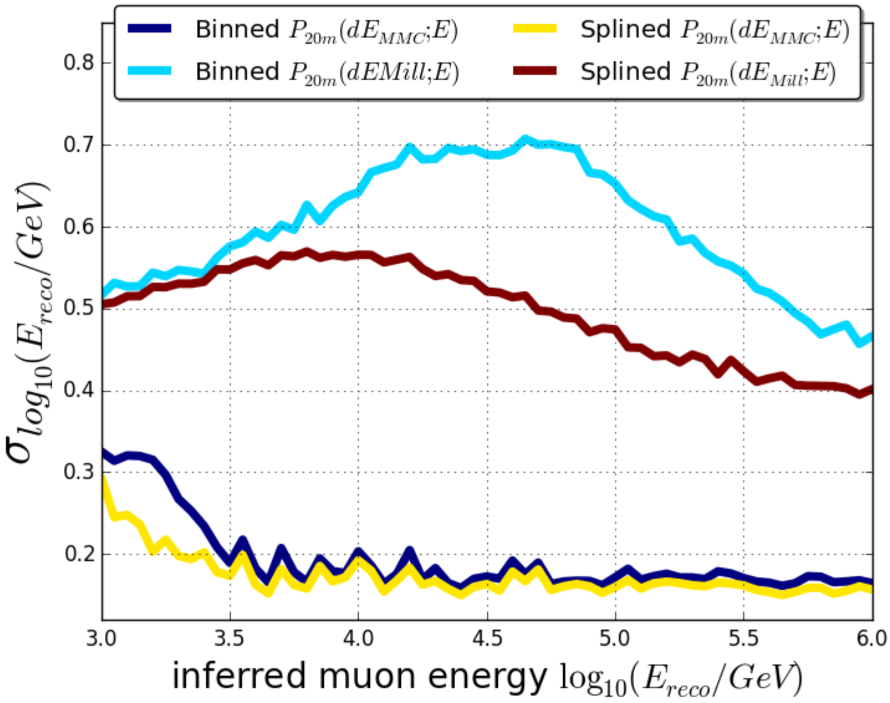


Figure C.7: The ALF reconstruction resolution for the Edepillim reconstruction on Millipede and MMC energy losses over a range of target true muon energies using binned PDFs and spline fitted(unbinned) PDFs. The spline fitted PDF results in a better resolution for the Edepillim reconstruction using both Millipede and true losses, with the most significant improvement for the reconstruction using Millipede losses.

Bibliography

[1] V. F. Hess, "Observations of the Penetrating Radiation on Seven Balloon Flights," *Phys. Z.*, vol. 13, pp. 1084–1091, 1912.

[2] T. K. Gaisser, "Cosmic Rays at the Knee," in *Energy Budget in the High Energy Universe* (K. Sato and J. Hisano, eds.), pp. 45–55, Mar. 2007.

[3] K. Kotera and A. V. Olinto, "The Astrophysics of Ultrahigh-Energy Cosmic Rays," *Annu. Rev. Astron. Astrophys.*, vol. 49, pp. 119–153, Sept. 2011.

[4] J. J. Beatty and S. Westerhoff, "The highest-energy cosmic rays," *Annual Review of Nuclear and Particle Science*, vol. 59, no. 1, pp. 319–345, 2009.

[5] A. M. Hillas, "The Origin of Ultra-High-Energy Cosmic Rays," *Annual Review of Astronomy and Astrophysics*, vol. 22, pp. 425–444, 1984.

[6] E. Fermi, "On the Origin of the Cosmic Radiation," *Phys. Rev.*, vol. 75, pp. 1169–1174, Apr 1949.

[7] C. Grupen, *Astroparticle Physics.* SpringerLink: Springer e-Books, Springer, 2005.

[8] T. Stanev, *High Energy Cosmic Rays.* Springer Praxis Books, Springer, 2010.

[9] M. G. Aartsen and et al., "Search for a diffuse flux of astrophysical muon neutrinos with the IceCube 59-string configuration," *Phys. Rev. D*, vol. 89, p. 062007, Mar. 2014.

[10] U. F. Katz and C. Spiering, "High-energy neutrino astrophysics: Status and perspectives," *Progress in Particle and Nuclear Physics*, vol. 67, pp. 651–704, July 2012.

[11] M. Longair, *High Energy Astrophysics.* Cambridge University Press, 2011.

[12] W. Pauli., “Open letter to the group of radioactive people at the gauverein meeting in tbingen,” 1930.

[13] F. L. Wilson, “Fermi’s Theory of Beta Decay,” *American Journal of Physics*, vol. 36, pp. 1150–1160, Dec. 1968.

[14] F. Reines, C. L. Cowan, F. B. Harrison, A. D. McGuire, and H. W. Kruse, “Detection of the free antineutrino,” *Phys. Rev.*, vol. 117, pp. 159–173, Jan 1960.

[15] G. Bellini, L. Ludhova, G. Ranucci, and F. L. Villante, “Neutrino oscillations,” *ArXiv e-prints*, Oct. 2013.

[16] J. K. Becker, “High-energy neutrinos in the context of multimessenger astrophysics,” vol. 458, pp. 173–246, Mar. 2008.

[17] S. N. Ahmed and et al., “Measurement of the Total Active ^{8}B Solar Neutrino Flux at the Sudbury Neutrino Observatory with Enhanced Neutral Current Sensitivity,” *Physical Review Letters*, vol. 92, p. 181301, May 2004.

[18] M. H. Ahn and et al., “Search for Electron Neutrino Appearance in a 250km Long-Baseline Experiment,” *Physical Review Letters*, vol. 93, p. 051801, July 2004.

[19] R. Gandhi, C. Quigg, M. Hall Reno, and I. Sarcevic, “Ultrahigh-energy neutrino interactions,” *Astroparticle Physics*, vol. 5, pp. 81–110, Aug. 1996.

[20] S. Schoenen, *Discovery and characterization of a diffuse astrophysical muon neutrino flux with the iceCube neutrino observatory.* Dissertation, RWTH Aachen University, Aachen, 2017. Verffentlicht auf dem Publikationsserver der RWTH Aachen University; Dissertation, RWTH Aachen University, 2017.

[21] M. G. Aartsen and et al., “Observation and characterization of a cosmic muon neutrino flux from the northern hemisphere using six years of icecube data,” *The Astrophysical Journal*, vol. 833, p. 3, Dec. 2016.

[22] J. Babson and et al, “Cosmic-ray muons in the deep ocean,” *Phys. Rev. D*, vol. 42, pp. 3613–3620, Dec 1990.

[23] I. B. et al, “The Baikal underwater neutrino telescope: Design, performance, and first results,” *Astroparticle Physics*, vol. 7, no. 3, pp. 263 – 282, 1997.

[24] E. e. a. Andres, “The AMANDA neutrino telescope: principle of operation and first results,” *Astroparticle Physics*, vol. 13, pp. 1–20, Mar. 2000.

[25] M. Ackermann and et al, “Optical properties of deep glacial ice at the South Pole,” *J. Geophys. Res. Atmos.*, vol. 111, no. D13, p. D13203, 2006.

[26] M. G. Aartsen and et al., “The IceCube Neutrino Observatory: instrumentation and online systems,” *Journal of Instrumentation*, vol. 12, p. P03012, Mar. 2017.

[27] A. Achterberg and et al., “First year performance of the IceCube neutrino telescope,” *Astroparticle Physics*, vol. 26, pp. 155–173, Oct. 2006.

[28] IceCube Collaboration, “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector,” *Science*, vol. 342, p. 1242856, Nov. 2013.

[29] M. G. Aartsen, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, M. Ahrens, D. Altmann, T. Anderson, C. Arguelles, T. C. Arlen, and et al., “Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data,” *Physical Review Letters*, vol. 113, p. 101101, Sept. 2014.

[30] M. G. Aartsen and et al., “Evidence for Astrophysical Muon Neutrinos from the Northern Sky with IceCube,” *Physical Review Letters*, vol. 115, p. 081102, Aug. 2015.

[31] M. Honda, T. Kajita, K. Kasahara, S. Midorikawa, and T. Sanuki, “Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data,” *Phys. Rev. D*, vol. 75, p. 043006, Feb 2007.

[32] R. Enberg, M. H. Reno, and I. Sarcevic, “Prompt neutrino fluxes from atmospheric charm,” *Physical Review D*, vol. 78, p. 043005, Aug. 2008.

[33] The IceCube Collaboration, M. G. Aartsen, and et al., “The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part II: Atmospheric and Astrophysical Diffuse Neutrino Searches of All Flavors,” *ArXiv e-prints*, Oct. 2015.

[34] R. Abbasi and et al., "An improved method for measuring muon energy using the truncated mean of de/dx," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 703, pp. 190 – 198, 2013.

[35] D. E. Groom, N. V. Mokhov, and S. I. Striganov, "Muon stopping power and range tables 10-MeV to 100-TeV," *Atom. Data Nucl. Data Tabl.*, vol. 78, pp. 183–356, 2001.

[36] J.-H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin, and J. B. Tjus, "Proposal: A tool for propagation of charged leptons," *Computer Physics Communications*, vol. 184, no. 9, pp. 2070 – 2090, 2013.

[37] H. Bethe and W. Heitler, "On the stopping of fast particles and on the creation of positive electrons," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 146, no. 856, pp. 83–112, 1934.

[38] R. Kokoulin, A. Petrukhin, and S. Kelner, "About cross section for high-energy muon bremsstrahlung," *Moscow Engineering Physics Inst*, 1995.

[39] H. A. Bethe and N. F. Mott, "The influence of screening on the creation and stopping of electrons," *Proceedings of the Cambridge Philosophical Society*, vol. 30, p. 524, 1934.

[40] R. P. Kokoulin and A. A. Petrukhin, "Influence of the nucleon form factor on the cross section of electron pair production by high energy muons," *Proceedings of the XII International Conference on Cosmic Rays*, vol. 6, 1971.

[41] L. B. Bezrukov and E. V. Bugaev, "Nucleon shadowing effects in photonuclear interactions," *Sov. J. Nucl. Phys.*, vol. 33, p. 635643, 1981.

[42] W. Lohmann, R. Kopp, and R. Voss, "Energy loss of muons in the energy range 1-10000 GeV," *CERN*, vol. 85, March 1985.

[43] P. Lipari and T. Stanev, "Propagation of multi-tev muons," *Phys. Rev. D*, vol. 44, pp. 3543–3554, Dec 1991.

[44] J. Beringer, J. F. Arguin, and et al., “Review of particle physics,” *Phys. Rev. D*, vol. 86, p. 010001, Jul 2012.

[45] M. G. Aartsen and et al., “Energy reconstruction methods in the IceCube neutrino telescope,” *Journal of Instrumentation*, vol. 9, p. 3009P, Mar. 2014.

[46] N. Whitehorn, J. van Santen, and S. Lafebre, “Penalized splines for smooth representation of high-dimensional Monte Carlo datasets,” *Computer Physics Communications*, vol. 184, pp. 2214–2220, Sept. 2013.

[47] R. P. Petrukhin and A. A. Kokoulin, “Theory of the Pair Meter for High Energy Muon Measurements,” 1987.

[48] W. Eadie, *Statistical Methods in Experimental Physics.* North-Holland Pub. Co., 1971.

[49] ANTARES Collaboration, “Transmission of light in deep sea water at the site of the ANTARES neutrino telescope,” *Astroparticle Physics*, vol. 23, pp. 131–155, Feb. 2005.

[50] C. Finley, “Weighting NeutrinoGenerator Events with OneWeight in IceSim 2.0,” *IceCube Internal Document*, Oct. 2007.

[51] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *European Physical Journal C*, vol. 71, p. 1554, Feb. 2011.

[52] J. van Santen and N. Whitehorn, “Photospline: smooth, semi-analytic interpolation for photonics tables,”

[53] P. H. C. Eilers and B. D. Marx, “Flexible smoothing with b -splines and penalties,” *Statist. Sci.*, vol. 11, pp. 89–121, 05 1996.

[54] C. de Boor, *A Practical Guide to Splines.* Applied Mathematical Sciences, Springer New York, 2001.